Volume 16, Number 1

Print ISSN: 1533-3604
Online ISSN: 1533-3590

JOURNAL OF ECONOMICS AND ECONOMIC EDUCATION RESEARCH

The *Journal of Economics and Economic Education Research* is owned and published by Jordan Whitney Enterprises, Inc. Editorial Content is controlled by the Allied Academies, a non-profit association of scholars, whose purpose is to support and encourage research and the sharing and exchange of ideas and insights throughout the world.

The *Journal of Economics and Economic Education Research* is owned and published by Jordan Whitney Enterprises, Inc., PO Box 1032, Weaverville, NC, 28787, USA. Those interested in communicating with the *Journal*, should contact the Executive Director of the Allied Academies at info@.alliedacademies.org.

EDITORIAL REVIEW BOARD

EDITORIAL REVIEW BOARD

TABLE OF CONTENTS

MARKET ATTRACTIVENESS EVALUATION OF SUB-SAHARAN AFRICA, APPLYING SWOT ANALYSIS AND AHP METHODS

Peter Symon Nganga, Nagasaki University
Yukihiro Maruyama, Nagasaki University

ABSTRACT

Purpose: *Market attractiveness of a nation is primarily driven by the macro factors Politics, Economy, Social and Technology (PEST). Generally, traditional market selection analysis emphasizes on purely macroeconomic and political factors which perceives the market attractiveness to be related only to these two sets of factors. At the outset the analysis fails to account for social/cultural dynamism which has great impact on future potential market attractiveness and the entire nation's economy. Therefore, the objective of this research is to expand the tradition model of economic & financial and political with social/ cultural issues, technology, & infrastructure and present a model which goes beyond the traditional PEST analysis to PESTI to evaluate the potential market attractiveness of 44 Sub-Saharan countries while prioritizing on social cultural issues.*

Design/Methodology/approach: *First, through literature review the author identifies the strengths, weaknesses opportunities and threats (SWOT) of the African continent. Secondly the author applies Analytical Hierarchy Process model to solve the complex multi-criteria decisions of evaluating the market attractiveness of a country with respect to global macro environment indicators. Statistical data from various credible sources were adopted for the weights calculation in AHP sub criteria level instead of the commonly applied intensity rating. Absolute measurement criteria are weighted independently of the evaluation of the alternative.*

Findings: *The resulting priorities revealed attractive market growth potential and sourcing opportunities in Mauritius that, might otherwise have been overlooked. Mauritius have integrated social cultural and politics with economic factors making it the best option among the 44 evaluated Sub-Saharan countries, various regional trading blocs are also a possibility. The analysis also helped us to gain a better understanding of the trade-offs in the decision making process and a clearer understanding of the effectiveness of AHP absolute measurements in multi criteria decision problem.*

Research Limitations/ Implications: *Follow up studies with more sub-criteria will contribute to the development of a better assessment model for the market attractiveness. The choice of criteria that are considered can vary depending upon the researcher or the industry which the businesses want to engage in therefore the model depends on personal preferences.*

Practical implication: *A useful tool or road map for Multinational National Enterprises searching for new markets in developing countries.*

Originality value: *The research develops the body of knowledge on market attractiveness by addressing the shortcomings of the traditional Macro analysis (PEST) and expands the three prior studies on developing countries market potential analysis. In addition, the author adopted normalized statistical data on the sub criteria instead of the commonly applied intensity rating.*

Paper Type: *Research Paper*

Key words: *Globalization, Market attractiveness, Sub-Saharan Africa, Analytical Hierarchy Process, SWOT, criteria, decision alternatives*

INTRODUCTION

Currently, convincing argument can be made that globalization has led to advancement in technologies and this advancement has ushered a new world without boundaries. Countries are currently interconnected by flows of knowledge and information through computers, TVs, and satellites and also the web and the internet have revolutionized and hastened how business is conducted globally. (Bhandari & Heshmati, 2005). Hence, ideas exchange has increased tremendously thus altering consumer perception and preferences while boosting countries' economies, trade, technologies and improved health care. However, understanding the global business environment and its complexities is challenging especially when, each nation's market environment is composed of unique cultural, political, legal and economic characteristics that defines or dictates how business is conducted in the nation's boundaries; this set of national characteristics may differ greatly from nation to nation. Consequently, globalization can be helpful for achieving genuine development but these conditions are not sufficient, and indeed, managed poorly, the local response or lack of response to globalization can ultimately do more harm than good for developing countries. (Tadaro & Smith, 2003). Regardless, this phenomena has opened up new markets that were once crossed to foreign companies, and its ultimate effect on trade will only increase the importance of standalone nations and regional trading blocs.

Although there are tons of literatures that have examined issues regarding to the selection of global emerging markets that are attractive to Multinational Enterprises (MNEs) in the developed countries. There are few academic studies that address the attractiveness of

developing countries such as in Latin America, Far East Asia, and Mediterranean Africa. However, currently, empirical research on opportunities in Sub-Saharan nations is very limited both in quantity and focus. Therefore, research expands the understanding of the potential of yet another untapped developing nations with great market attractiveness potential. Generally, traditional market selection analysis relies on purely macroeconomic and political factors at the outset of the analysis, and fails to account for a developing market's vitality and future potential resulting from rapid change, national attributes that affect specific sectors and market receptiveness. (Sakarya & Eckman & Hyllegard, 2006). Moreover, the analyzed available information for those developing countries primarily deals with economics and economic systems, which perceives the market attractiveness to be related to only two sets of factors deriving from two points of view: economic & financial and political. (Saaty 1980; Tripodo & Dazzi, 1995; Abid & Bahlouh, 2011). However, these two set of factors are not sufficient to fully analyze the complexities of market attractiveness in Sub-Saharan African nations, which differs from those of developing nations in Mediterranean Africa, Asia and Latin America in terms of social cultural, political systems the level of economic development, and geographic climatic conditions, their markets are characterized by higher degree of risk than their developed counterparts. Therefore, this research incorporates the two sets of factors deriving from two points of view: economic & financial and political and expands them with macro indicators namely, social cultural, technological factors and infrastructure in the multi criteria decision problem analysis., into few easily understandable priorities based on the degree of conformity between potential or existing market environmental factors at the macro level (national level) emphasizing on the external environment of each country`s social/cultural, political/ legal, economic and technological applying AHP methodology. The research presents important opportunities for MNEs operating in the areas of financial services, mineral resources, infrastructure development, industrial sector development and various other related consultancies.

GENERAL PURPOSE

Traditionally, market attractiveness of a country is primarily driven by the macro factors Politics, Economy, Social and Technology (PEST). Therefore, in this research market attractiveness of the (nations) is considered to be consistent and robust growth of economic and non-economic factors at the macro level in recent years. Host market potential is considered one of the most important explanatory factors in country attractiveness and the market selection and constitutes a primary driver in a firms venture into international markets. (Yoshida, 1987). In this case the author explains, a country potential could be related to a set

of variables economics /financial and political legal, social cultural, and technology/infrastructure which have on going improvement for the business environment, exponential growth in trade and investment and of substantial improvements in the quality of human life. Due to the complexities of the Sub- Saharan political economy emphasis should be more on social/cultural factors being the major contributor of civil strife in Africa. The regions multi-ethnic composition at times causes tribal conflicts which affects economic growth. Evaluation of the market attractiveness in Sub-Saharan countries focus should be both on standalone and regional bloc attractiveness. Some standalone attractive markets also happen to be globally strategic markets and this is the arena where the current and future global competition occurs. (Gillespie et.al, 2007). Moreover, most Sub-Saharan countries are landlocked which offers them geographic proximities with identical climatic conditions.

The goals of this undertaking are to contribute to the field of research and to the business sphere in the following:

(a) Expert knowledge which incorporates economic and non-economic factors for sound judgment arising from various sources for MNEs Senior Managers while upgrading the traditional model (PEST) to the new (PESTI)

(b) Incorporate conventional relative measurements with conventional absolute measurements on AHP methodology for multi-criteria decision making in the global environment

(c) Suggest and highlight new markets opportunities in Sub Saharan Africa.

(d) Stimulate academic discussion about the various problems and opportunities in Sub-Saharan Africa.

SWOT ANALYSIS SUB SAHARAN AFRICAN REGION:

SWOT analysis is a simple widely used qualitative tool which examines a company, industry or a nation's strengths and weakness (internal factors) with opportunities and threats (external factors).The analysis provides the basic outline in which to perform analysis of decision situations. In this situation the author uses the tool to examine the strengths, weaknesses, opportunities and threats of the Sub-Saharan African continent as a block market. The strength of the African continent as a whole is its richness in natural resources; it has 50 % of the world's gold, most of the world's diamonds and chromium, 90 % of the cobalt, 40% of the world's potential hydroelectric power, 65 % of the manganese, millions of acres of untilled arable farmland as well as other natural resources. (Williams, 1997). Moreover, last year, five of the top 10 fastest growing economies globally were Sub-Saharan African countries: South Sudan, Sierra Leone, Gambia, Mozambique and the Democratic Republic of Congo. There was no doubt that, these real GDP growth figures were impressive as economies

were growing rapidly from a low base. Some were even down to one-off factors such as in the case of South Sudan where 2013 real GDP growth of 32.1% (the highest in the world) was a forecast rebound from a severe contraction of 53.0% in 2012 when the government shut down oil production. However, the region overall is expected to maintain the second fastest economic growth globally, with a forecasted real GDP growth of 47.1% in 2013-2020 (Eghbal, 2013). Therefore, MNEs senior managers should focus on targeted and tailored strategies for each country overall, the size of the Sub-Saharan African.

Last year, a markets attractiveness survey conducted by Ernest &Young, overall ranked Africa fifth out of nine other regions, ahead of the former Soviet states, Eastern Europe, the Middle East, Western Europe and Central America, the respondents ranked Africa as a more attractive place for investments a significant improvement from the survey conducted in the year 2011 which Africa was slightly ahead of Soviet states and Central America. (Ernst &Young`s, 2013). The opportunities in Africa are increasingly evident, by the year 2035, the continent will have the largest workforce with over half of the population currently under the age of 20; over the last decade improvements in macroeconomics and a burgeoning and fast growing South-South trade and investment flow (with over US$170 billion with China alone). Across various sectors Africa presents ample prospects with US$2.6trillion of revenue expected by 2020 across resources, agriculture, consumer and infrastructure, of which US$1.4 trillion will be solely in consumer industries. Furthermore, being the primary recipients of international development aid these countries' presents crucial opportunities for those organizations operating in areas such as infrastructure development, industrial sector development, financial services, and other consultancies.

Despite the considerable improvement over the past decade, the major threat is, extreme poverty still remains widespread in the region. Over 800; million people are still struggling against extreme poverty and the situation may worsen with the population projected to be 1.7 billion by 2050. (JICA, 2013). According to the latest global poverty update for the first time since 1981, less than half of the African population of 47% lived below $ 1.25 a day in 2008, the rate was 51% in 1981. However, the $1.25 a day poverty rate in Sub-Saharan Africa has fallen 10 percentage points since 1999. The facts are, severe development challenges still remains in Africa, where approximately one in every two people currently, lives on $1.25 a day (World Bank, 2012) the level of poverty is also aggravated by the Neglected Tropical Diseases (NTDS) and HIV virus. Since the epidemic of HIV aids globally about 70 million people are affected by the HIV virus and 35 million have died from it, Sub-Saharan Africa still remains the most severely affected region with nearly every 20 adults (4.9%) living with HIV

and accounting for 69 % of the people suffering with HIV worldwide. (World Health Organization, 2013).

Far from the HIV aids virus, Neglected Tropical Diseases are not necessarily serious health hazards; nonetheless they are an integral cause of poverty to many families. NTDs are primarily found in Asia, the Pacific, Central and South America. However, the majority of the people infected with NTDs live in Sub-Saharan Africa and in order to achieve the millennium goal of poverty eradication there is a greater need for NTDs control and if this can be attained it will be a huge relief on developing countries vulnerable economies. (O`Brien MP, 2008). Tribal conflicts and terrorism are other major problems, though the rate of tribal wars occurrence has subsided considerably still the problem crops up now and then due to uneven distribution of wealth from natural resources, and cattle grazing and watering pasture areas. Although terrorism is a relatively new problem brought by religion differences, the Islamist radicals have taken advantage of weak central governments, un manned porous borders, under-trained and under-paid police forces and flourishing drug cartels, for example, early this year there was shootings at an upscale shopping mall in Kenya instigated by a Somali militant group, again, early this year, an Islamic Maghreb conquered Mali (Olga Khazan, 2013). On May 20th of this year, a bomb blasted at Central Nigeria claiming over a hundred lives. Relatively, the negative image of the continent as a whole conceals the complex diversity of the economic performance and the existence of investment opportunities in individual nations. In spite of all these obstacles, economists expect to see US$ 1.4 trillion in spending by African consumers in 2020. (Mahinda, 2013).

In historical and geographical perspective the entire continent consists of 54 small independent nations in total, 48 out of those considered as the Sub-Saharan region. Unfortunately Europe's arbitrary post-colonial demarcation left Africans bunched into countries that don't represent their heritage; a contradiction that still troubles the region even today. The artificial borders have often led to border conflicts. May be if the borders demarcation had no uncertainties' between Eritrea and Ethiopia, Mali and Burkina Faso, Nigeria and Cameroon, Senegal and Mauritania those wars could have been avoidable. (Zartman, 2001). One most common characteristic, they are landlocked therefore, supply chain requires frequent border crossings which is very difficult to manage due to poor infrastructures maintenance and lack of modern technology. This is a major hindrance for economic growth since it imposes excessive extra cost on transportation of goods. Solving these barriers may reduce the cost in supply chain as well as promote trade across industries, while attaining regional social and economic integration. Since the year 2005, the Japanese International Corporation Agency has studied and supported the development of cross-border transport

infrastructure (CBTI), in their recent research study, CBTI has defined infrastructure as the requirement for transportation that crosses multiple national borders, and the infrastructure that comprehensively includes physical "hard Infrastructure" such as ports, railroads, highways, cargo transshipment facilities, national border facilities, weighbridges (truck scales), and inland container depots (ICDs), as well as "soft Infrastructure" such as cross-border transport laws, regulations related to border crossing (e.g. Customs clearance, quarantine), and organizational systems and resources for smoothly operating and maintaining the hard infrastructure mentioned above (JICA, 2013).

According to the regional director of the African Development Bank, Africans have known for more than 50 years that, the infrastructures lags behind and it should be prioritized due to the fact that, the African growth has caused huge demographic shift from rural to the urban areas and the infrastructure has not kept pace with the growth. The director emphasizes that, over 30 countries have prolonged power problem, and transportation cost are on the rise increasing the cost of goods by approximately 75% in some of the landlocked nations. He emphasizes that, for the next decade, Africa needs to spend almost US$90 billion a year to upgrade and maintain its crumbling infrastructure. (Faal, 2013).

APPLYING AHP MODEL, MARKET ATTRACTIVENESS:

There are various Multi-criteria decision analysis methodologies and Multivariate statistical methodologies used as geometric representations supporting multi-criteria decision making but, at times it's hard to interpret the final result as a map of the environment due to its dependence on the measurable statistical properties of the data rather than on, more correctly, the perception of the problem and its political and social ramifications as they apply to each nation. Therefore, AHP developed by (Saaty, 1980), appeared to be the most appropriate suited model being a simple decision analysis model which combines subjective judgment and system approach. Currently, AHP model is widely used to solve various problems in Africa, such as, the suitability of community based management approach in forest reserves of Rwanda, (Masozerra & Alavalapatib & Jacobsonc & Shresthab, 2006), assessment for potential multi-airport system in Cape Town South Africa, (Zietman & Vanderschureen, 2014) and for screening urban transport projects in Accra Ghana, (Jones & Tefe & Opuku, 2013). The method applicability to multifaceted, multi-period and unrestricted problems made it the most preferred model for developing countries market attractiveness potential.

The application of AHP usually takes the following three basic steps, structuring the hierarchy, setting priorities and maintaining rational consistency. In this research, structuring the hierarchy, the decomposition of the overall goal is to find the market attractiveness

(potential) in Sub-Saharan African countries. The top level of the hierarchy refers to the goal which in this case is “market attractiveness or potential”. The subsequent levels include the elements that affect the decision (criteria or attributes), in this case Economic, Political/legal, Social/Cultural, Technology and Infrastructure these are the four main Macro-factors that most substantially influence a nation’s attractiveness. The second level includes elements (sub-criteria’s) that contribute to the definition of the first level criteria, in this case the author have prioritized factors that affects social cultural issues most. The bottom level consists of the decision alternatives (44 Sub Saharan countries, there was insufficient data for meaning analysis on four countries).

Setting the priorities entailed determining, for each level of the hierarchy, the relative importance between each pair of factors. The pairwise judgment started from the second level attributes (Economic, Political/Legal, Social/Cultural and Technology) to the lowest (Alternatives-44 Sub-Saharan Nations). However, the original Saaty`s AHP used relative measurements and had shortfalls in which it cannot deal with a situation involving many alternatives such as the 44 Sub-Saharan Nations, to overcome such predicament Saaty proposed an absolute measurement which we have incorporated with the dominant alternatives method proposed by (Kinoshita & Nakanishi, 1997) a new type of Analytical Hierarchy Process designed to deal with cases in which the weight of the criteria valley in accordance with the alternative chosen as the dominant view (Conventional absolute measurement). Finally the country is evaluated on its performance with respect to each sub-criterion using normalized data for the most appropriate rating grade. Normalization was necessary in this case since indicators such as GDP and inflation impacts the model differently, for example, highest GDP is best but a higher inflation is bad. The results were then weighted and combined to yield weights with respects to the major sub-criteria’s and the ranking of potential markets (nations) is the synthesized results or the output sought.

Criteria’s: Usually, the screening process of nations markets starts with gathering relevant information on each nations and screening out those un- desirable. The first stage involves applying macro- indicators (PEST) to discriminate between nations that present basic opportunities and those that either offers little or no opportunity or higher risk. Traditionally the Macro-Indicators describes the total market in terms, of PEST, emphasis primarily on political and economic attributes however, for the purpose of in-depth analysis the author has expanded the traditional (PEST) and included Infrastructure attribute to make (PESTI) in the methodology (AHP) with emphasis on social/cultural, **Sub-Criteria:** The second level includes elements (sub-criteria’s) that contribute to the definition of the first level criteria the

four main macro-indicators. For example, GDP, inflation contributes to economic criteria, Global peace Index, and CPIA contributes to political/legal criteria e.tc.

Alternatives: Each decision alternative (44 Sub-Saharan nations) contributes to each criterion in a unique way. Applying AHP the author specifies a mathematical process to synthesize the information on relative importance of the criteria and the preferences for the decision alternatives to provide an overall priority ranking of the decision alternatives. In the market attractiveness evaluation problem AHP will use the Authors personal preferences to provide a priority ranking of the forty four nations in terms of how well each nation meets the overall objective of being the best nation with the most appeal. Moreover, in the evaluation of the elements in the sub criteria level statistical data was adopted instead of the intensity ratings.

Establishing priorities using AHP: Pairwise comparison of the four criteria for the market attractiveness. In each of the above comparisons the author selected the most important criterion and then expressed sound judgment on how much more important the criterion is. Table 1 provides a summary of the five pairwise comparisons provided for the attractive market selection problem. However, the author notes that, the flexibility of AHP can accommodate the unique preferences of each individual researcher or analyst. The choice of criteria that are considered can vary depending upon the researcher or the industry which the businesses want to engage in. AHP methodology can accommodate any set of criteria depending with the decision maker.

(Figure 1.) (HIERARCHY FOR THE MARKET SELECTION PROBLEM

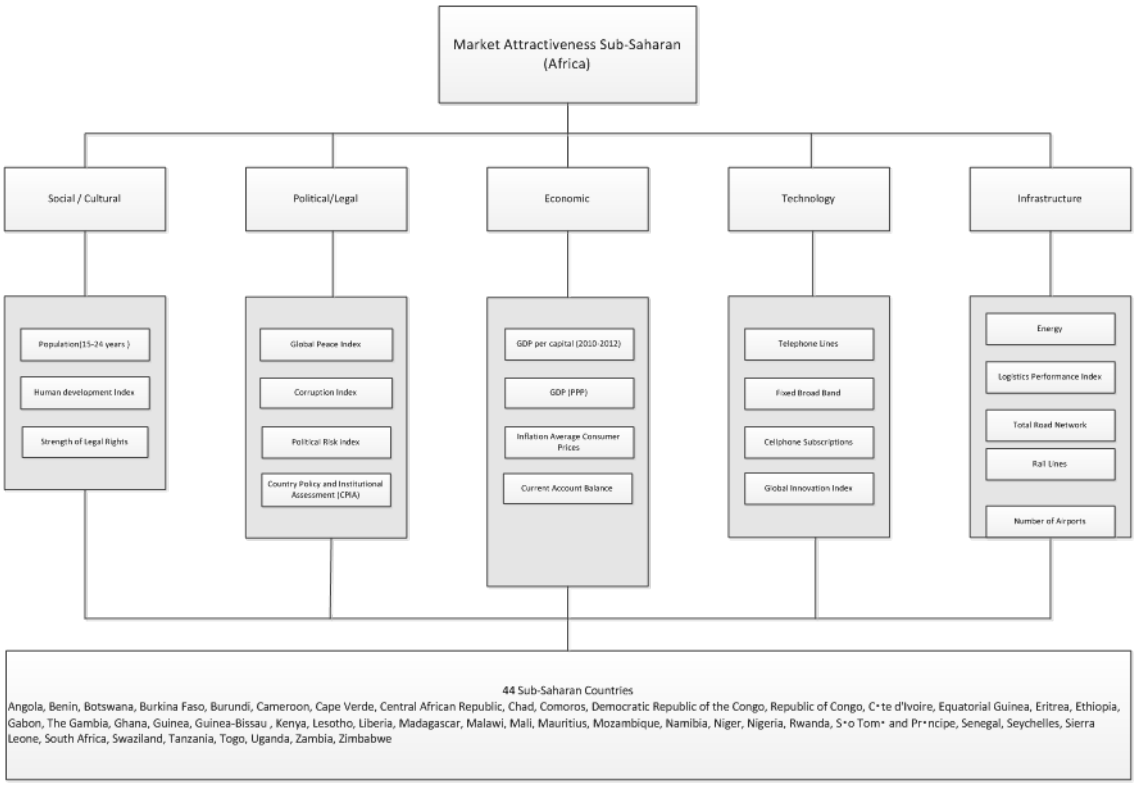


Table 1 is the summary of the five pairwise comparison Social/Cultural and politics are equally important and together these two criteria dominates the remaining criteria.

Table 1 Macro Indicators Comparisor (Criteria)

	Social/Cul	Politics	Economic	Technology	Infrastructure	Weight Priority
Social/Cul	1	1	3	6	3	0.3491
Politics	1	1	3	6	3	0.3491
Economic	1/3	1/3	1	3	1	0.1262
Technology	1/6	1/6	1/3	1	1/3	0.0495
Infrastructure	1/3	1/3	1	3	1	0.1262

In table 2 below, is the second level which includes elements (sub-criteria) that contribute to the definition of the first level criteria or the five main macro-indicators. Population is the dominant indicator.

Table 2. Social/Cult Comparison Matrix (Sub-Criteria)				
	population15-24	Human Developm	Strength of lega	Weight Priority
population15-24	1	3	3	0.6000
Human Developm	1/3	1	1	0.2000
Strength of legal r	1/3	1	1	0.2000

In table 3 below, the dominant indicator is the global peace.

Table 3. Political/Legal Comparison Matrix (Sub-Criteria)					
	Global Peace Index	Corruption Index	Political Index	Country Policy	Weight Priority
Global Peace Inde	1	3	3	3	0.5000
Corruption Index	1/3	1	1	1	0.1667
Political Risk Inde	1/3	1	1	1	0.1667
Country Policy &	1/3	1	1	1	0.1667

In table 4 below, the author identified GDP per Capital, GDP (PPP), and Current Account as the dominant criteria all three are equally important.

Table 4. Economics Comparison Matrix (Sub-Criteria)					
	GDPPer capital	GDP(PPP)	Inflation Rate	Current Acoun	Weight Priority
GDP Per Capital	1	1	1/3	1	0.1667
GDP (PPP)	1	1	1/3	1	0.1667
Inflation rate	3	3	1	3	0.5000
Current Accoint E	1	1	1/3	1	0.1667

In table 5 below, telephone lines and global innovations dominates the rest of the criteria.

Table 5. Technology/ Comparison Matrix (Sub-Criteria)					
	Telephone lines	Fixed Broad Ban	Cellphone Subs	Global Innovat	Weight Priority
Telephone lines	1	3	3	1	0.3750
Fixed Broad Ban	1/3	1	1	1/3	0.1250
Cellphone Subs	1/3	1	1	1/3	0.1250
Global Innovation	1	3	3	1	0.3750

In table 6 below, the dominant indicators are the Energy consumption, logistics index, and the total road network.

Table 6. Infrastructure Comparison Matrix (Sub-Criteria)

	Energy Consumption	Logistics Index	Total Road Ne	Rail Lines	Airports	Weight Priority
Energy Consump	1	1	1	2	3	0.2601
Logistics Index	1	1	1	2	3	0.2601
Total Road Netw	1	1	1	2	3	0.2601
Rail Lines	0.5	0.5	0.5	1	2	0.1378
Airports	1/3	1/3	1/3	0.5	1	0.0819

Formula applied: Below is the formula applied for the Criteria, Sub criteria's and the alternatives (Countries) To conclude the final result the alternatives two measurements were applied conventional relative measurements and conventional Absolute measurement The criteria are weighted independently of the evaluation of the alternative. MATLAB was also used to derive the final result of the Country`s ranking

$$W_1^T = (w_1, w_2, w_3, w_4)$$

$$W_2^T = (w_{e1}, w_{e2}, w_{e3}, w_{e4})$$

$$W_3^T = (w_{p1}, w_{p2}, w_{p3}, w_{p4})$$

$$W_4^T = (w_{s1}, w_{s2}, w_{s3})$$

$$W_5^T = (w_{t1}, w_{t2}, w_{t3}, w_{t4})$$

$$\begin{pmatrix} w_1 * w_{e1} \\ w_1 * w_{e2} \\ w_1 * w_{e3} \\ w_1 * w_{e4} \\ w_2 * w_{p1} \\ w_2 * w_{p2} \\ w_2 * w_{p3} \\ w_2 * w_{p4} \\ w_3 * w_{s1} \\ w_3 * w_{s2} \\ w_3 * w_{s3} \\ w_4 * w_{t1} \\ w_4 * w_{t2} \\ w_4 * w_{t3} \\ w_4 * w_{t4} \end{pmatrix}$$

$$S = \begin{pmatrix} s_{1\,1} & \cdots & s_{1\,15} \\ \vdots & \ddots & \vdots \\ s_{44\,1} & \cdots & s_{44\,15} \end{pmatrix}$$

$$E = S_{ij} * W^T$$

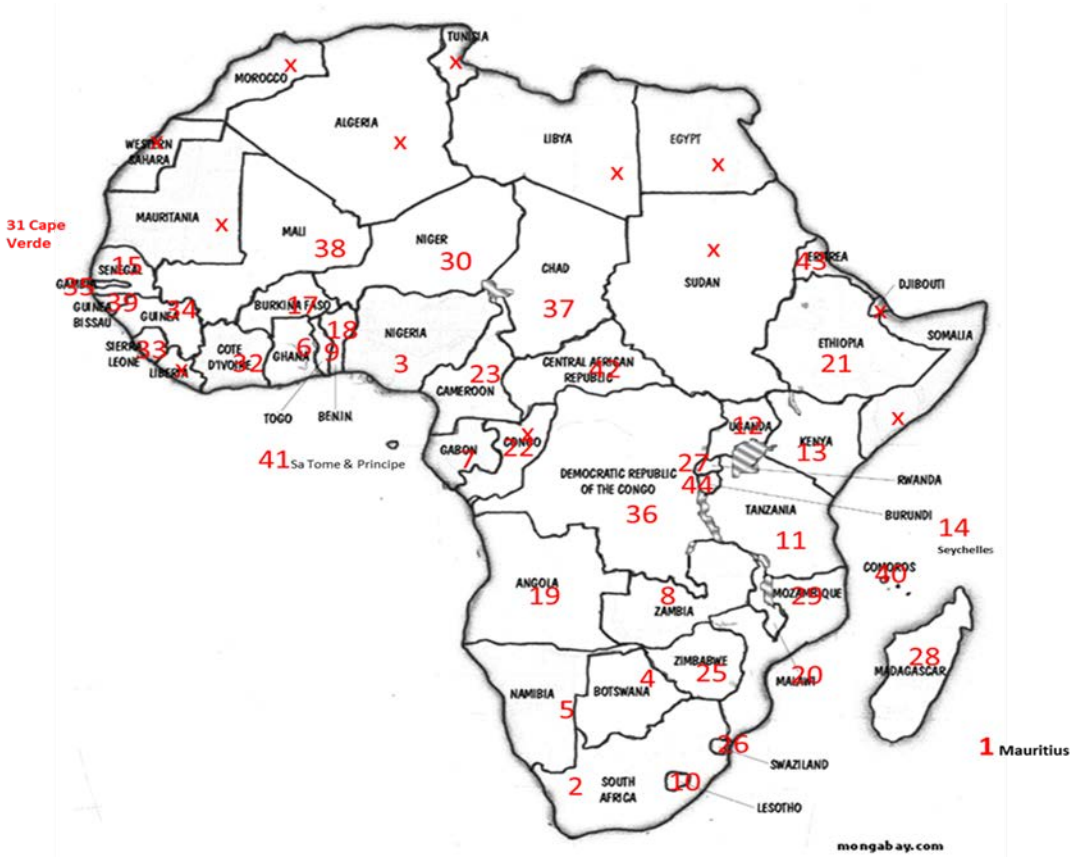
All sub-criteria and total weights

Sub-Criteria's	Weights
population15-24	0.2094
Human Development Index	0.0698
Strength of legal rights	0.0698
Global Peace Index	0.1745
Corruption Index	0.0582
Political Index	0.0582
CPIA	0.0582
GDP Per Capital	0.0210
GDP (PPP)	0.0210
Inflation rate	0.0631
Current Account Bal	0.0210
Telephone lines	0.0186
Fixed Broad Band	0.0062
Cellphone Subs	0.0062
Global Innovations index	0.0186
Energy Consumption	0.0328
Logistics Index	0.0328
Total Road Network	0.0328
Rail Lines	0.0174
Airports	0.0103

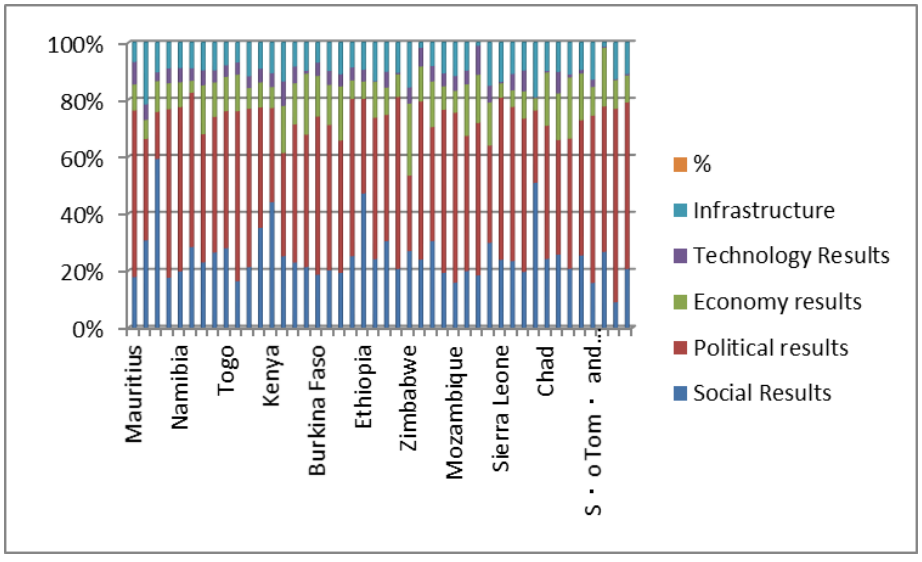
AHP ranking of the decision alternatives total weights

Ranking	Coutry Name	Total	Ranking	Coutry Name	Total	Ranking	Coutry Name	Total	Ranking	Coutry Name	Total
1	Mauritius	0.5416	12	Uganda	0.3579	23	Cameroon	0.2980	34	Guinea	0.2422
2	South Africa	0.5186	13	Kenya	0.3569	24	Liberia	0.2924	35	The Gambia	0.2404
3	Nigeria	0.4962	14	Seychelles	0.3480	25	Zimbabwe	0.2892	36	Democratic Republic of the Cong	0.2337
4	Botswana	0.4770	15	Senegal	0.3322	26	Swaziland	0.2889	37	Chad	0.2306
5	Namibia	0.4504	16	Equatorial Guinea	0.3315	27	Rwanda	0.2790	38	Mali	0.2290
6	Ghana	0.4277	17	Burkina Faso	0.3304	28	Madagascar	0.2784	39	Guinea-Bissau	0.2190
7	Gabon	0.3912	18	Benin	0.3270	29	Mozambique	0.2757	40	Comoros	0.2094
8	Zambia	0.3742	19	Angola	0.3263	30	Niger	0.2694	41	S・o Tom・ and Pr・ncipe	0.1963
9	Togo	0.3695	20	Malawi	0.3199	31	Cape Verde	0.2644	42	Central African Republic	0.1787
10	Lesotho	0.3665	21	Ethiopia	0.3193	32	C・te d'Ivoire	0.2621	43	Eritrea	0.1644
11	Tanzania	0.3589	22	Republic of Congo	0.2995	33	Sierra Leone	0.2469	44	Burundi	0.1423

Figure 2 Respective countries map ranking



The map indicates visually there are tendencies within the Southern regional, East African and West African regional bloc markets. May be this could be attributed to favorable geographical locations easily accessible to the ports.



The graph above shows the heat-wave of the newly designed PESTI in percentages on how each issue contributes to the final decisional making analysis.

CONCLUSION & RECOMMENDATIONS

Generally, firms prefers to venture in attractive markets that are graded higher in attractiveness with low risk, high profitability and where competitive advantage is attainable however, attaining all those mentioned factors in a globalized market environment is not a simple task. It requires various well augmented strategies to venture even into those nations classified as a high risk. Conventional wisdom may suggest that, MNEs might postpone entry of the developing markets however, some types of first –mover advantages may be higher in these economies. (Arnold & Quelch, 1998). Therefore, it is necessary for MNEs from developed nations to enter into these markets in developing countries with the proper entry and exit strategy configurations to attain, existing market expansion, strategic resource seeking, and natural resource seeking and host country's location advantages. Entry, exist strategy involves various considerations though the importance of these considerations varies by industry and by the main objective of each company.

This research addressed the need for more expansion on the traditional PEST to PESTI. The author proposed additional criteria (infrastructure) which includes energy, logistics, and communications infrastructures. However, the proposed criteria are not substitutes for existing market attractiveness evaluation method, but are intended to expand and address the shortcomings of traditional PEST. Due to the complexities of political economy and the social structure in Sub-Saharan Africa emphasis or priority should be on social/cultural issues, in the evaluation model. The multi-ethnic composition at times causes tribal conflicts which affects the entire economy. Therefore, economic growth impacted by the tribal conflicts cannot be exactly revealed by economic & financial and political criteria alone rather these criteria tend conceal the potential markets attractiveness of the region. Moreover, focus should be both on standalone attractiveness and regional bloc attractiveness. Some standalone attractive markets also happen to be globally strategic markets and this is the arena where the current and future global competition occurs. (Gillespie et.al, 2007).

The author combined SWOT analysis and AHP model to determine respectively the relative importance of the criteria and alternatives as pertaining to the social/cultural, political, economic, technology and infrastructure positions of Sub-Saharan African countries and the result provides basis for MNEs and other international businesses to make decisions regarding the potential market attractiveness of Sub-Saharan Africa. SWOT analysis highlighted the strengths and the hidden opportunities of the Sub-Saharan Africa while the potential threats

and weaknesses may help senior managers in risk hedge management. Applying AHP model, the author also analyzed and calculated percentages of the each macro indicator and its contributions to the overall goal of the countries market attractiveness. The AHP priorities results indicated that, Social cultural and political issues had the greatest impact on market attractiveness in Sub-Saharan African market attractiveness. The resulting priorities revealed attractive market growth potential and sourcing opportunities in Mauritius that, might otherwise have been overlooked applying the traditional PEST model. The government of Mauritius have integrated social cultural, infrastructure and politics with economic factors making it the best option among the 44 evaluated Sub-Saharan countries followed by South Africa and Nigeria respectively. Since 1980s, the government of Mauritius undertook social and economic reforms breaking down barriers to improve the ease of doing business. For example, personal and income tax were halved, empowered labor laws, directed lending policy and banks were obliged to led to the export processing zones at lower rates than anywhere else in Africa. (Hon Xavier Luc Duval the Vice Prime Minister of Mauritius). Thus the importance of Mauritius aligning the social/cultural and political issues with economic development may have played a central role in the final analysis criterion in the author's decision making. Various regional trading blocs were also a possibility, the Southern region as a bloc market had the best potential among other regions.

These analysis also helped us to gain a better understanding of the trade-offs in the decision making process and a clearer understanding of the effectiveness of AHP absolute measurements in multi criteria decision problem while combining both theory and practicality . However, more research is required with countries in the map with priorities between 0.2995 and 0.2190 perhaps by adding more factors to the sub-criteria level would offer the decision makers a chance to modify the weights and note the resulting map.

REFERENCES

Amit.K. Bhandari, Almas Heshmati, (April 2005). Measurement of globalization and its variations among countries and region overtime. MTT Economic Research, Seoul National University and IZA Bonn.

Angela Edmunds, Anne Morris, (2000). The problem of information overload in business organizations: a review of literature International Journal of Information Management 20 (2000) 17-28

Arnold, DJ. Quelch,J.A.(1998). New strategies in emerging markets, Sloan Management Review, Vol.40,pp. 7-20

Central Intelligence The world Fact Book. www.cia.gov/library/publications/the-world-factbook/fields/2053.html accessed Nov-26, 2013

Davina Zietsman, Marianne Vanderschuren (2014). Analytic Hierarchy Process assessment for potential multi-airport systems -The case of Cape Town. Journal of Air Transport Management.

Diana Layfield , Africa CEO Standard Chartered (2013). Seeking opportunity managing risk Ernst&Young`s attractiveness survey Africa 2013 Getting down to business.www.comattractiveness

Dr. Denise Youngblood Coleman, Editor in Chief, CountryWatch Inc. www.countrywatch.com

Ebrima Faal, (2013). Regional Director, African Development Bank The Infrastructure deficit and opportunities in Africa Ernst&Young`s attractiveness survey Africa 2013 Getting down to business.www.comattractiveness

Eizo Kinoshita & Masatake Nakanishi, (June 1999). Proposal of new AHP Model in light of dominant relationship among alternatives. Journal or Operation Research Society of Japan Vol.42, No.2,

Fathi Abid,& Slah Bahloul, (2010). Selected MENA COUNTRIES` ATTRACTIVENESS TO G7 INVESTORS,Working Paper No. 531 Economic Research Forum (ERF) 7 Boulus Hanna Street Dokki, Cairo Egypt.

G.TRIPODO & N. DAZZI (1995). Market Attractiveness of Developing countries, Socio-Econ. Plann. Sci. Vol.29, Elsevier Science LTD.

Gerald Mahinda, Managing Director BrandHouse (2013). The rise of African Consumer: Myth or reality. Ernst&Young`s attractiveness survey Africa 2013 Getting down to business. business.www.comattractiveness

Kate Gillespie & Jean-Pierre Jeanet & H. David Hennessey (2007). Global Marketing, Houghton Mifflin Company, 222 Berkeley Street, Boston, MA 02116-3764

http://www.transparency.org/research/cpi/overview

I. William Zartman, (May 2001). The Clear marking of boundaries between African nations can help prevent costly and debilitating border conflicts. Bordering on War Foreign Policy.

IC Publications, LTD. (April 2010). FDI into Africa The iron is hot Afri Bus no363, & ColdBath SQ, London ECIR, 4LQ England.

Japan firms look to tap high growth." The Nikkei Weekly (Japan). (May 27, 2013): 998 words. LexisNexis Academic. Web. Date Accessed: 2013/11/25.

Kei Yoshizawa (June 2013) Achieving Economic Transformation for Inclusive and Sustained Growth in Africa: Prospect and Challenges Japan International Cooperation Agency (JICA).

Levis R. Cabrera & German E. Giraldo (2009). A Multiple Criteria Decision Analysis for the FDI in Latin American Countries. Department of Industrial Engineering, University of Puerto Rico, Mayaguez, Puerto Rico 00680, USA.

Media Eghbal Country Insight Managing Editor (April 2013). Why Africa's' Economic Fundamentals are Ripe for an Investment Boom. Euromonitor International.

Melly, Paul. African telecoms to witness boom: the continent is one of the fastest growing markets for mobile providers. MEED Middle East Economic Digest (26 Apr. 2013):s26. Business Insights: Global Web, 17 Nov. 2013

Michael P. Todaro & Stephen C. Smith (2003) Economic Development Eighth Edition, Addison Wesley.

Professor Walter E. Williams (June 4, 1997) Economic Freedom and progress Economics, George Mason University

Michel K. Masozeraa & Janaki R.R. Alavalapatib &Susan K. Jacobsonc& Ram K. Shresthab, (2006). Assessing the suitability of community-based management for the Nyungwe Forest Reserve, Rwanda. Forest Policy and Economics 8 (2006) 206– 216

Olga Khazan, (2013). Why Africa Is the New Terrorism Hub, Retrieved May 21, 2014, from http://www.theatlantic.com/international/archive/2013/09/why-africa-is-the-new-terrorism-hub/279956/

Richard M. Steers & Luciara Nardon (2006). Managing in the Global Economy, M.E. Sharpe, INc., 80 Business Park Drive. Armonk, New York 10504

Rozann W. Saaty (2003). The Analytical Hierarchy Process (AHP) for Decision Making and The Analytic Network Process (ANP) For Decision Making with Dependence and Feedback. Creative Decision Foundation 4922 Ellsworth Avenue Pittsburgh, PA 15213.

Sema Sakarya, &Molly Eckman & Karen H. Hyllegard, (2006). Market selection for international expansion: Assessing opportunities in emerging markets, International Marketing Review, Vol. 24 Iss: 2, pp.208 – 238

Stanley A.J. & Clipsham P.S. (1997). Information overload-myth or reality? IEE Colloquium Digest No. 97/340

Stephen O`Brien MP "The Neglected Tropical Diseases, A Challenge we could rise to-will we?" Report for the All-Party Parliamentary Group on Malaria and Neglected Tropical Diseases" 9House of Commons APPMG 2008/9

Steven Jackson World Bank Group (April 2012). Poverty Reduction ad Economic Management in Sub-Saharan Africa (AFR PREM)

Steven Jones a, Moses Tefe, Seth Appiah-Opoku, (2013). Proposed framework for sustainability screening of urban transport projects in developing countries: A case study of Accra, Ghana. Transportation Research Part A: Policy and Practice An International Journal

Thomas L. Saaty & Luis G. Vargas (2001). Models, Methods, Concepts & Applications of the Analytic Hierarchy Process"Springer Science + Business Media New York 1st Edition.

TRANSPARENCY INTERNATIONAL the global coalition against corruption,

World Economic "Forum Global Information Technology Report 2013" www.weforum.org/ /

World Economic and Financial Surveys World Economic Outlook databasehttp://www.imf.org/external/pubs/ft/weo/2013/02/weodata/index.aspx

World Health Organization (2013), "Adult HIV Prevalence (15-49 years), 2011 By WHO region

Yoshida, M. (1987). Macro-micro analyses of Japanese manufacturing investment in the United States, Management International Review, Vol.27 No.4, pp.250-56.

SHOW ME THE MONEY: EVALUATING THE IMPACT OF AN ONLINE PROFESSIONAL DEVELOPMENT COURSE FOR MIDDLE AND SECONDARY CLASSROOM TEACHERS

Kimberly P. Clayton-Code, Northern Kentucky University

ABSTRACT

An interactive, online professional development course titled, "Show Me the Money: Personal Finance for Middle and Secondary Classrooms" was recently developed as part of a collaboration between a regional metropolitan university and a state economic education organization. The course was designed to encourage middle and secondary educators to integrate personal finance concepts across the curriculum using materials from the Council for Economic Education resources and activities, and focused on teaching strategies to differentiate learning for all students. This case study presents the initial results of the collaborative professional development course in terms of teacher attitude towards economic education and student learning outcomes. Initial analysis of the results indicates a positive increase in teacher attitudes as well as a statistically significant increase in student knowledge of personal finance concepts.

SHOW ME THE MONEY: EVALUATING THE IMPACT OF AN ONLINE PROFESSIONAL DEVELOPMENT COURSE FOR MIDDLE AND SECONDARY CLASSROOM TEACHERS

This has been an amazing course and I really find it hard to be able to express which resources are going to be most useful to me. I have thoroughly enjoyed each of them, even though I tend to like the interactive websites, games, and quizzes due to the attractiveness of this type of resource for students. I am posting my webliography and can't wait to see others' and add them to mine. I have seen such an increase in my students' understanding of the importance of personal finance just in these 8 weeks.

A recent collaboration between a state economic education advocacy organization and metropolitan university resulted in the development and facilitation of an interactive, online professional development course, "Show Me the Money: Personal Finance for Middle and Secondary Classrooms". The course was designed to encourage middle and secondary educators to integrate personal finance across the curriculum using materials from the Council for Economic Education resources and activities, and focused on teaching strategies to differentiate learning for all students. This case study presents the initial results of the collaborative professional development course in terms of teacher content knowledge, attitude towards economic education, and student learning outcomes.

RESEARCH OVERVIEW

Consumers who function successfully in today's economy are confronted with economic and personal finance decisions on a daily basis. These decisions have become increasingly difficult due to the numerous goods and services available in the market and the free exchange of information via quickly changing technologies and the World Wide Web. Today's students need to learn how to make sound decisions by accessing and assessing consumer information and comparing and evaluating goods and services. The President's Advisory Council on Financial Literacy defines financial literacy as the "ability to use knowledge and skills to manage financial resources effectively for a lifetime of financial well-being" (2008 Annual Report to the President). Financial literacy includes skills like long-term vision and planning for the future along with the discipline to use those skills on a daily basis. Middle and secondary classrooms must include appropriate opportunities for students to be engaged in personal finance related activities that prepare them to make sound decisions throughout their lives.

A number of studies have shown the financial knowledge of today's middle and secondary students to be deficient (Lusardi, Mitchell & Curto, 2010; Loibl, 2010; and Spielhofer, Kerr, and Gardiner, 2010). A recent JumpStart Coalition survey indicated that the financial literacy of high school students had fallen to its lowest level ever. Furthermore, the Survey indicated that that 75 percent of young American adults are likely to lack the skills needed to make beneficial financial decisions. For example, only 16.8 percent of high school seniors and 19.2 percent of college students thought that stocks are likely to have higher average returns than savings bonds, savings accounts and checking accounts *over an 18 year period*. Only 27.3 percent of high school seniors realized that interest on a savings account is taxable if one's income is high enough. Only about 40 percent of high school seniors realize that their own health insurance could stop if their parents become unemployed (Mandell, 2008).

"Just as it was not possible to live in an industrialized society without print literacy—the ability to read and write—so it is not possible to live in today's world without being financially literate. To fully participate in society today, financial literacy is critical" (Council for Economic Education, 2011). The Council for Economic Education publishes a biennial report detailing the status of economic and financial education in the United States. The most recent report (2011) found that only 14 states required a course in personal finance to be offered. The report also found that less than 20 percent of teachers felt competent to teach personal finance topics.

In a report from Gutter, Copur, and Garrison (2010), the authors found that states where a financial education course was required had the highest reported financial knowledge and were more likely to display positive financial behaviors and dispositions. These students were more likely to save, less likely to max out credit cards, less likely to make late payments, more likely to pay off credit cards in full each month, less likely to be compulsive buyers, and more likely to be willing to take average financial risk.

In addition to students' lack of personal financial knowledge, middle and secondary teachers have also indicated their unease with teaching personal finance and economic concepts. Way and Holden (2009) found that 89 percent of K-12 teachers agree that students should take a financial education course or pass a competency test for personal finance before graduating from high school, relatively few teachers believe they are adequately prepared to teach such topics. Their research also found that teachers were open to the idea of integrating personal finance into their curriculum, but needed assistance with integrating the concepts into their discipline.

Walstad (2001) and Walstad and Rebeck (2000) found that infusing economics and personal finance into existing content courses do not provide effective economics education. In addition, social studies teachers may not have adequate training in economics and personal finance leading to the need for specialized training in order to equip teachers with knowledge to integrate economics and personal finance into their teaching.

The good news is that there is a body of work showing that student learning increases as the economic training of their teachers increase (Watts, 2006; Schober, 1984; and Staubs, 2008). In a pilot program by Cargill, Jurosky, and Wendel (2008), the researchers found that skilled teachers were able to achieve significant improvements by working with formally trained economists in a less-than-full-college-course environment. Their approach included a brief teacher training coupled with integrated discussions of economic content and teaching strategies, and the teachers were provided funding to design and implement modules themselves.

In terms of the types of training preferred by teachers, Becker, Greene, and Rosen (1990) found that teachers prefer pragmatic training that includes strategies and lesson plans for introducing content to their students. It was in response to this body of research, that the *Show Me the Money* online professional development course was developed to determine if a six-week online professional development course for teachers could have an impact not only on teachers' understanding of personal finance concepts, but more importantly on the knowledge of their students.

PROFESSIONAL DEVELOPMENT COURSE

The "Show Me The Money: Personal Finance for Middle and Secondary Classrooms" online professional development course was designed to encourage middle and secondary educators to integrate personal finance across the curriculum using materials from the National Council for Economic Education resources and activities, and focused on teaching strategies to differentiate learning for all students. (The course development and assessment were conducted with the financial support from a Council on Economic Education Excellence in Economic Education grant.)

The course was developed as a partnership between a state economic education advocacy organization, a metropolitan university, and the state department of education's e-learning branch. For this project, the state economic education advocacy organization provided the content specialists for the course design, a team of reviewers, development of the course, resources to support the course, collection of the student assessments, and evaluation of the effectiveness of the online professional development course. The state department of education's e-learning branch's mission is to assist educators by "providing high quality, facilitated, media-rich and interactive online professional development and training thus helping to improve P-12 education" (e-Learning Kentucky, 2009). This department provided training for the course design team, the team of reviewers, a course facilitator, registration and hosting of the course, surveys of the teachers, and resources to supplement the online professional development course. The online professional development courses offered were available online 24 hours a day, 7 days a week, allowing the participating teachers access to the course at their convenience. Finally, the researcher served as the course designer and facilitator as well as conducted the evaluation of the course assessments.

The course explored personal finance concepts included in JumpStart Coalition's National Standards for Personal Finance Education and state curriculum content standards

through interactive, hands-on lessons, activities, and resources that could be immediately integrated into participating teachers' middle and secondary classroom instruction. The goals for the course included the opportunity for participating teachers to:

1. *Explore developmentally appropriate personal finance concepts and principles for middle and high school students;*
2. *Examine personal finance content standards at both state and national levels;*
3. *Learn about theory and best practice in teaching personal finance concepts to middle and high school students;*
4. *Explore and analyze web sites and other technology for use in learning personal finance content and develop a list of technology resources appropriate for teaching personal finance content to middle and high school students;*
5. *Create instruction that will incorporate technology and will facilitate students' learning of personal finance concepts;*
6. *and Reflect on classroom practices.*
7. *As a final product participants will develop and share an integrated, interactive webliography that can be integrated into their classroom instruction.*

The structure of the course itself consisted of six sessions—an orientation session and six content sessions. The sessions were offered online through the Blackboard course management system. Each session lasted for one week, beginning on a Wednesday and ending on the following Tuesday. During each session, each participating teacher completed an assignment that included readings, activities, and online discussions. Surveys were also required at the beginning and upon completion of the course. The estimated time for completing each weekly session or module was 2-4 hours. The outline for the individual sessions was as follows:

Orientation	During the Orientation Session, you will have an opportunity to explore the course website, to experiment with the course tools and to introduce yourself to your facilitator and fellow participants. Orientation takes place during the first week.
Session 1	Financial Responsibility and Decision Making—This session introduces ways that students can apply reliable information and use economic decision making resulting in personal finance decisions. Participants will begin to explore resources for integrating personal finance into their classroom instruction. They will also begin developing a plan that includes resources and how they might integrate them into their own classroom.
Session 2	Income and Careers—This session explores ways that middle and high school students can use a career plan to develop their personal income potential. Lessons presented in this session will explore career options, identify sources of income and examine factors that affect one's income and career.
Session 3	Planning and Management—This session presents ways in which students can organize their personal finances and use a budget to manage cash flow. Ideas and lessons for developing a person financial plan, charitable giving, and systems for keeping and using financial records will be presented.
Session 4	Credit and Debt—This session will explore ways in which one can maintain creditworthiness, borrow at favorable terms, and manage debt. Ideas for presenting the costs and benefits of various types of credit, developing a strong credit history, ways to avoid debt problems, and consumer credit laws will be presented. In addition, lessons that focus on the topic of identity theft will be provided.
Session 5	Saving and Investing—This session presents the case for developing middle and high school students' understanding of saving and investing in terms of implementing a diversified investment strategy that is compatible with personal goals. Participants' personal finance concepts webliography and plans for integrating personal finance concepts into their own classrooms should be finalized and final surveys will also be completed.

As a final product, participants created an economics and personal finance webliography. This webliography was a compilation of website resources that were woven throughout the course. The idea was for the participating teachers to leave the course with multiple online resources for their teaching. For each concept listed, the teachers researched and found three resources that could be used to teach the concept. The template included the name of the website, the URL, a description of how the website could be used in the classroom setting, and the personal finance concepts taught.

In order to successfully complete the course, the following criteria were used for evaluating successful participation in and completion of the course:

Course Activities	Participants were expected to complete the required course readings and activities as posted in each of the session assignment pages. Participants were expected to post reflections about the assigned readings and the completed activities in the online course discussion.
Discussion Postings	Participants were expected to respond to the discussion prompt(s) in each of the sessions with an original posting that includes information from readings and activities. Participants were expected to respond thoughtfully to the postings of other course participants in each session, possibly referencing readings, etc. Guidelines for original discussion postings and responses were available in the course orientation.
Participant Assessment	Participants were expected to complete the final product and post their work as directed during the final course session. Participants completing team products were expected to work collaboratively with their team and include the names of all team members on the final product posting within the course.
Course Surveys	Participants were expected to complete all course surveys within one week of the assignment.

Each participating teacher completed three surveys—a demographic survey, a course pre/post assessment, and an overall course evaluation. They also conducted pre and post-test assessments with the students in their individual classrooms. The purpose of these assessments was to evaluate the effectiveness of the course in terms of the impact on teacher knowledge and integration of personal finance concepts into classroom teaching. In addition, the goal of the assessment was to evaluate the impact of the teacher participation in the online professional development economics course on student knowledge of personal finance concepts.

EVIDENCE OF DOCUMENTED STUDENT ACHIEVEMENT CHANGES

Three thousand, one hundred, twenty-four students from 56 middle and secondary classrooms participated in this initial assessment of the *Show Me the Money* online professional development course. The student data was analyzed using a t-test with a pooled estimate of population standard deviation to determine the significance of the difference between the mean pretest and posttest scores. The t-test was used to determine if the difference in the sample means for knowledge (pretest scores and posttest scores) can be attributed to the influence of the

teachers' participation in the *Show Me the Money* online professional development course, or if the difference could have happened by chance. Because the pre- and post-test sample sizes varied, the t-test statistic was computed using a pooled estimate of population standard deviation. The following t-statistic which compensates for uneven samples was used to test the hypothesis about two means:

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (u_1 - u_2)}{s^* \sqrt{\frac{1}{n_1} - \frac{1}{n_2}}}, \quad \text{where } s^* = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}},$$

where X-bar is the pre/post-test mean, u_1 is the pretest population mean, u_2 is the post-test population mean, s_1 is the pretest variance, s_2 is the post-test variance, n_1 is the pretest population size, and n_2 is the post-test population size. The n for the pre-and post tests were different as some classes experienced growth whereas others experience attrition. This test was used to test the hypothesis that there was no significant difference between the pre- and post-test student means for each question on the instrument.

The results of the t-test analyses presented in Table 1 show that the middle and secondary students scored significantly higher on all 25 of the 25 questions related to knowledge of personal finance concepts. In fact, the student scores for all of the questions with the exception of one (question #15) were statistically significant at the 99% level.

Table 1: Statistical Analysis of Student Assessment

	Question 1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
PreMean	0.19	0.82	063	0.81	0.76	0.37	0.66	0.37	0.55	0.66
PostMean	0.58	0.91	0.81	0.84	0.87	0.46	0.73	0.59	0.75	0.82
T-Test	-39.83	-14.22	-17.27	-3.45	-12.63	-6.59	-6.03	-16.93	-16.03	-15.48
Sig.	**	**	**	**	**	**	**	**	**	**

	Question 11	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
PreMean	0.63	0.34	0.32	0.60	0.22	0.26	0.53	0.53	0.48	0.13
PostMean	0.77	0.56	0.47	0.75	0.25	0.42	0.66	0.66	0.53	0.29
T-Test	-12.80	-13.56	-12.85	-13.12	-1.88	-15.05	-10.51	-10.87	-3.08	-19.56
Sig.	**	**	**	**	*	**	**	**	**	**

	Question 21	**22**	**23**	**24**	**25**
PreMean	0.28	0.21	0.35	0.45	0.51
PostMean	0.41	0.42	0.41	0.64	0.73
T-Test	-11.61	-20.07	-4.09	-14.44	-18.24
Sig.	**	**	**	**	**

EVIDENCE OF TEACHER ATTITUDINAL CHANGES

Each of the 56 participating teachers completed course surveys to evaluate the impact of the course on their attitudes toward and knowledge of economic concepts. In terms of demographic information, the participating teachers all taught in middle and secondary schools throughout the state and included grades five through twelve. In terms of their teaching experience, 8% had taught between one to four years, 34% had taught between five to ten years, 50% had taught between 11-20 years and 8% had taught 21 years or more. Their educational levels included 8% with at least a bachelor's degree, 42% with master's degrees, and 50% with Rank I degrees. (A Rank I degree is a graduate degree consisting of at least 30 graduate credit hours focused on a particular area of study.)

The course teacher survey instruments included attitudinal and demographic questions. Questions number five through twelve were designed using a bi-polar adjective format which is a variant of the semantic differential test design. These eight questions addressed teacher attitude towards personal finance and attitudes toward personal finance teaching and curriculum materials.

Because the pre- and post-test teacher sample sizes varied, the t-test statistic was computed using a pooled estimate of population standard deviation. This test was used to test the hypothesis that there was no significant difference between the pre- and post-course means for each question on the instrument. The results of the statistical analysis are shown in Table 2.

Table 2: Teacher Attitudinal Questions

Question	Mean Score Pre/ Post-Training
1. I would describe my attitude towards personal finance as being _____. VERY FAVORABLE 1 --- 2 --- 3 --- 4 --- 5 --- 6 --- 7 VERY UNFAVORABLE	pre: 2.00 post: 1.47
4. To the best of my knowledge, quality curriculum materials for teaching personal finance are _____ . READILY AVAILABLE 1 --- 2 --- 3 --- 4 --- 5 --- 6 --- 7 DIFFICULT TO FIND	pre: 3.12 post: 1.53
5. I am _____ with internet or web site locations regarding personal finance. UNFAMILIAR 1 --- 2 --- 3 --- 4 --- 5 --- 6 --- 7 VERY FAMILIAR	pre: 2.71 post: 4.65
6. At the current time, I feel _____ my ability to find quality curriculum materials on personal finance. CONFIDENT OF 1 -- 2 -- 3 -- 4 -- 5 -- 6 -- 7 GENERALLY UNCOMFORTABLE WITH	pre: 3.00 post: 1.41
7. I feel that I am currently _____ to teach personal finance curriculum. HIGHLY QUALIFIED 1 -- 2 -- 3 -- 4 -- 5 -- 6 -- 7 GENERALLY UNQUALIFIED	pre: 3.00 post: 1.65
8. Given my current training, I _____ teaching a unit(s) on economics. DREAD 1 --- 2 --- 3 --- 4 --- 5 --- 6 --- 7 LOOK FORWARD TO	pre: 3.53 post: 4.82

The results for the attitudinal questions, shown in Table 2, focused on teacher perceptions of personal finance, teachers' perceptions of the availability of materials, curriculum and websites to teach personal finance, and teachers' perceptions of their abilities to teach personal finance. The responses to question one were very favorable initially—so much so that significant improvement may have been difficult to achieve. Even so, the mean scores

improved, although the improvement was not statistically significant. Overall, the t-test statistic indicated that—as a result of the professional development training—teachers felt significantly more confident about their ability to deliver financial education. Specifically, the scores indicate that teachers felt much more confident about the availability, and of their efforts to find personal finance education materials. Additionally, the teachers felt more qualified to teach personal finance curricula, and that they looked forward to teaching a unit on personal finance concepts.

Specific anecdotal feedback from teachers that further supported their positive attitudes as a result of the course. This course feedback included comments such as:

> *"I have learned so much from this course! I have gathered so many valuable economic resources to use with my students and to share with my colleagues! Thanks a million!"*

> *"I have learned so much from this class. … I have gathered all the websites and books and lessons that have been posted and I'm going to share them with the teachers at my school. They are always grateful to have resources sent to them. I feel a lot more comfortable teaching personal finance, especially since I have never taught it before! I know now websites to find ideas and I'm sending the webliography links to my media specialist so she can post them to our website. Thanks for a great class."*

Other anecdotal feedback included: 100% of the teachers indicated that they found the discussion board topics to be beneficial; 71% rated the overall quality of the workshop to be excellent with 29% rating it very good; 100% found the webliography to be an effective and usable project; 100% found the content of the workshop to be easily transferable to the classroom; and 94% felt knowledgeable about teaching personal finance in their own classrooms as a result of participating in the course.

Most importantly however, the results of the student assessments indicate that the online professional development course did, in fact, improve middle and secondary students' knowledge of key personal finance concepts—which was the preferred outcome of the course. Students who understand basic personal finance concepts are more able to make sound everyday financial decisions throughout their lives.

CONCLUSIONS

The use of professional development courses to assist teachers in their instruction of personal finance concepts to middle and secondary students is not a new concept. However, the advent of online instruction has presented a new opportunity for teachers to develop their pedagogical and content knowledge in the content area of personal finance. This study investigated the effects of a six week online professional development course, *Show Me the Money*, targeted specifically to middle and secondary teachers for the purpose of enabling them to increase their students' knowledge and understanding of basic personal finance concepts. Participating teachers completed pre-course and post-course assessment surveys and conducted pretest and posttest assessments of basic personal finance concepts with their students. The results of the initial offering of the online professional development course offer a positive, affirming impact of the course on teachers and most importantly on the content knowledge of their students.

REFERENCES

Becker, W., W. Greene, & S. Rosen. (1990). Research on high school economic education. *American Economic Review* 80 (May): 14-22.

Cargill, T.F., J. Jurosky, & J. Wendel. (2008). Implementing economics standards: A pilot transition program. *Journal of Economic Education* 39 (Spring): 126-134.

Council for Economic Education. (2011). Survey of the states: Economic and personal finance education in our Nation's Schools 2011. Retrieved June 30, 2012, from http://www.councilforeconed.org/wp/wp-content/uploads/2011/11/2011-Survey-of-the-States.pdf.

Gutter, M., Copur, Z., & Garrison, S. (2010). Financial capabilities of college students from states with varying financial education policies. Denver, CO: National Endowment for Financial Education.

Goetz, J., B.J. Cude, R.B. Nielson, S. Chatterjee, & Y. Mimura. (2011). College-based personal finance education: Student interest in three delivery methods. *Journal of Financial Counseling and Planning* 22(1): 27-42.

Harris Interactive Inc. (2012), The 2012 consumer financial literacy survey. The National Foundation for Credit Counseling and The Network Branded Prepaid Card Association. Retrieved June 22, 2012, from http://www.nfcc.org/newsroom/FinancialLiteracy/files2012/FLS2012FINALREPORT0402late.pdf.

Loibl, C. (2010). A look inside: Teacher information search for personal finance instruction. *International Journal of Consumer Studies* 34: 287-297.

Lusardi, A., O.S. Mitchell, & V. Curto. (2010). Financial literacy among the young. *The Journal of Consumer Affairs* 44 (2): 358-380.

Mandell, L. (2008). The financial literacy of young American adults: Results of the 2008 national Jump$tart Coalition survey of high school seniors and college students. Jump$tart Coalition.

Odom, M.O. (2008). *Economics education: Attitudes and practices of elementary and middle grade teacher candidates*. Dissertation Abstracts International Section A: Humanities and Social Sciences, Vol 68(10-A). University of Alabama.

Schober, H.M. (1984). The effects of inservice training on participating teachers and students in their economic classes. *Journal of Economic Education* 15 (4): 282-295.

Spielhofer, T., D. Kerr, & C. Gardiner. (2010). Personal finance education: Effective practice guide for schools. National Foundation for Educational Research.

Walstad, W. B. (2001). The status of economics in the high school curriculum. *Journal of Economic Perspectives* 15(Summer): 195-210.

Walstad, W.B. (2000). The status of economics in the high school curriculum. *Journal of Economic Education* 31 (Winter): 95-102.

Way, W.L. & K.C. Holden. (2009). Teachers' background and capacity to teach personal finance. National Endowment for Financial Education. Retrieved July 5, 2012, from: http://www.afcpe.org/assets/pdf/vol20_2way_holden.pdf.

APPENDIX : SHOW ME THE MONEY STUDENT ASSESSMENT

1. Which is true about most millionaires?
   a) They inherit most of their wealth.
   b) They work more than 40 hours a week.
   c) They work in glamorous jobs like sports or entertainment.
   d) They avoid investing in the stock market due to its riskiness.

2. Which strategy is most likely to improve most people's financial situation over a lifetime?
   a) using credit to spend more than a person earns
   b) making financial decisions quickly based on intuition
   c) saving early in life when a person begins earning an income
   d) gaining work experience early instead of continuing in school

3. If Jessica chooses to buy more auto insurance this year than she bought last year, then Jessica has probably decided that:
   a) the benefits of more auto insurance are greater than the costs.
   b) more auto insurance will reduce her risk of an accident.
   c) this added insurance will help her savings grow faster.
   d) she will drive more safely in the future.

4. Which of the following is typically true?
   a) Correct choices have no costs.
   b) People do not respond to incentives.
   c) Voluntary trade causes winners and losers.
   d) People's choices have consequences for the future.

5. Some people choose to stop attending high school before graduation. What is the opportunity cost to the individual of dropping out of high school?
   a) The lower income earned in the job market.
   b) The lower interest rates one is charged for credit.
   c) The higher taxes one pays for leaving school early.
   d) The payments made to companies hiring drop-outs.

6. The first three steps of economic decision-making are to:
   a) draw a conclusion, construct a model, and make a generalization.
   b) identify assumptions, make a policy, evaluate the policy.
   c) gather facts, form a theory, and conduct a simulation.
   d) define the problem, list alternatives, and state criteria.

7. Which is most likely to improve a person’s chances of getting a job?
   a) supplying a neat and accurate application
   b) preparing a resume only if an employer requests one
   c) avoiding classified job advertisements in newspapers
   d) dressing for an interview in a manner that is provocative

8. An approach to finding out about job opportunities that consists of contact with friends and family members who might be able to connect you to a potential employer is called
   a) logrolling.
   b) recruiting.
   c) connecting.
   d) networking.

9. Which question are employers allowed to ask job applicants during a job interview?
   a) Are you married?
   b) Do you attend a place of worship?

c) What do you feel are your weaknesses as they pertain to the job?
d) Do you have any disabilities that might affect your job performance?

10. A typical characteristic of entrepreneurs is that they
a) like to work part-time.
b) are willing to take risks.
c) would rather save than invest.
d) enjoy working under supervision.

11. The U.S. Bureau of Labor Statistics reveals that the fastest growing jobs will require which level of educational training?
a) an associate's degree or higher
b) postsecondary vocational training
c) no more than a high school diploma
d) work experience in a related occupation

12. What is the difference between gross and net pay?
a) Net pay is gross pay minus saving.
b) Gross pay is net pay minus saving.
c) Gross pay is net pay minus deductions.
d) Net pay is gross pay minus deductions.

13. Emily works full-time at Handy Dandy Hardware. Who pays the FICA (Social Security) contribution on Emily's wages?
a) Emily only
b) her employer only
c) Emily and her employer
d) the federal government

14. An employee works for 40 hours per week at $10 per hour. The deductions per week are $50 in federal income taxes, $20 in state income taxes, and $30 in Social Security taxes. What is the net pay for the week?
a) $300
b) $330
c) $350
d) $400

15. What is the opportunity cost of letting your interest compound in your savings account instead of withdrawing the interest as it is earned?
a) lower taxes in the current year
b) increased risk of losing savings
c) less money for current purchases
d) more interest on the savings account

16. Beginning to save while you are young is recommended by financial experts because it:
a) is easier to save when you first begin earning income.
b) is hard to save later in life when you have more income to spend.
c) allows you to lock in higher interest rates when you buy on credit.
d) lets compound interest work in your favor by earning interest on interest.

17. What is the general relationship between risk and reward?
a) the higher the risk, the lower the potential reward
b) the higher the risk, the higher the potential reward
c) the amount of risk does not influence potential reward
d) there is a relationship, but it is uncertain

18. A key advantage of getting credit is that it can help people:
    a) sell assets
    b) increase their net worth.
    c) reduce risk when investing in stocks.
    d) buy a good or service today and pay for it later.

19. Who generally benefits from a loan transaction?
    a) the lender only
    b) the borrower only
    c) both the borrower and the lender
    d) neither the borrower nor the lender

20. Which three things do creditors consider to be most important when judging a person's creditworthiness to buy a house or car?
    a) marital status, gender, location
    b) character, collateral, and capacity
    c) length of loan, credibility, comissions
    d) occupation, connections, income sources

21. What does a credit bureau do?
    a) extends credit to qualified buyers
    b) provides advice on how to use credit
    c) tracks the bill-paying habits of consumers
    d) sends warnings to people in credit trouble

22. Which type of financial institution typically charges the highest interest rates for loans?
    a) credit unions
    b) commercial banks
    c) savings and loans
    d) payday loan companies

23. Disposable income is the money that is:
    a) deducted from your paycheck.
    b) budgeted for variable expenses.
    c) spent or saved after deductions.
    d) saved and invested each month.

24. "Pay Yourself First" means that:
    a) all bills get paid before any saving.
    b) money is set aside for savings before spending.
    c) fixed expenses are paid before flexible expenses.
    d) credit cards pay for what you don't have as income.

25. Using a debit card to purchase a good is most similar to using a
    a) loan.
    b) check.
    c) credit card.
    d) money market account.

RETHINKING THE USE OF CONCEPT MAPS IN INTRODUCTORY ECONOMICS COURSES

Mariya Burdina, University Of Central Oklahoma

ABSTRACT

A Conceptual Chapter Map is a modified version of a traditional concept map, which in addition to illustrating the links among concepts, also provides explanations of concepts, which makes it a more practical learning tool since students can use it not only to understand the concepts, but also while studying or working on problem sets. This paper discusses the benefits of Conceptual Chapter Maps as compared to traditional concept maps and lecture notes, discusses why instructors of economics courses should consider conceptual chapter maps as a valuable learning tool, provides practical strategies for creating and delivering conceptual chapter maps, and reports students' attitudes toward the use of these maps in introductory economic courses.

INTRODUCTION

Concept maps are widely used in various educational disciplines to help students develop an understanding of a subject matter. Traditional concept maps show how concepts are related to each other by illustrating the links among multiple concepts. This paper will introduce a modified version of the traditional concept map, which I refer to as a Conceptual Chapter Map. The Conceptual Chapter Map not only shows the links among concepts, but also provides an explanation of each, thereby creating a more efficient learning process.

The idea of graphical concept organization was first proposed by David Ausebel in 1960, and further developed by Joseph Novak in the 1970s. Novak's version of concept organization is what educators now consider a traditional concept map: a tree diagram that shows the links among concepts. Research shows that concept maps significantly improve student learning and allow students to retain newly-obtained information more efficiently when compared to traditional lectures (Horton et al., 1993; Willerman and Macharg 1991). Traditional concept maps make the learning process easier and help students separate the most important information from what is not essential (Ellis, 2001). Concept maps are also used by instructors to promote successful collaborative learning in- and outside the classroom (Budd 2004; Baitz 2009).

Concept maps are widely used in various disciplines such as mathematics, physics, biology, and chemistry (Chiou 2008), but are seldom used in economics. By my personal observations, economics instructors are reluctant to use concept maps. This is unfortunate since this teaching tool can help students who are new to economics unravel the complexities of economic concepts.

In introductory economics courses, students develop a fundamental understanding of basic economic principles. It is during these classes that students learn economics "from scratch." Some may become overwhelmed with the amount of new information presented in class, for most topics students have several pages of notes, and it may be difficult for them to see how different concepts are related to each other. Traditional concept maps can help students organize their knowledge by creating a visual representation of how the concepts are linked, and

have shown to improve student learning (Jacobs-Lawson & Hershey 2002, Novak & Cañas 2008). Traditional concept maps are most helpful in developing the understanding of links between the concepts. However, on their own, these maps are not sufficient in helping students solve problems, since students have to refer to a text book or lecture notes to find explanations of the concepts being studied. Traditional lectures, on the other hand, provide students with definitions, explanations, and examples, but fail to show how concepts are related to one another. Traditional lectures and traditional concept maps complement each other, but are not very practical when used separately.

Research suggests that students learn better and with less difficulty when all the information they need is provided within a single source (Mayer 2009). The Conceptual Chapter Map as proposed in this paper follows the idea that a single information source benefits student learning. It combines the elements of traditional lecture notes with traditional concept maps, thereby providing students with a simple, visual, and easy-to-follow learning tool.

For example, while a traditional concept map may illustrate that Gross Domestic Product can be either real or nominal, a Conceptual Chapter Map takes this one step further and illustrates, using formulas, that Real GDP is calculated using base-year prices and Nominal GDP is calculated using current prices. The Conceptual Chapter Map shows student not only that there are several ways of calculating GDP, and explains the differences among these approaches. The following figures illustrate the differences between a Conceptual Chapter Map and a traditional concept map, both explaining concepts related to Gross Domestic Product. Figure 1 shows a traditional concept map, adopted from Johnston, Carol, University of Melbourne, and Figure 2 shows a Conceptual Chapter Map which I have used while teaching a Principles of Macroeconomics course during the Spring 2012 semester at the University of Central Oklahoma.

While both maps are relatively similar in the way they structure and illustrate the concepts, there are more explanations included in the Conceptual Chapter Map than in the traditional concept map, thus allowing students to answer most questions related to the GDP topic.

Figure 1. GDP: Traditional Concept Map

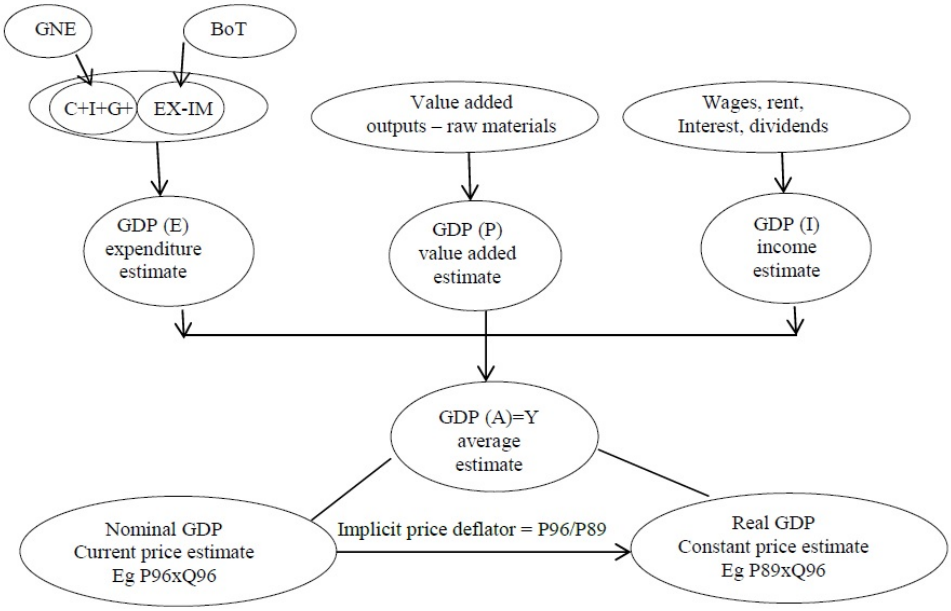
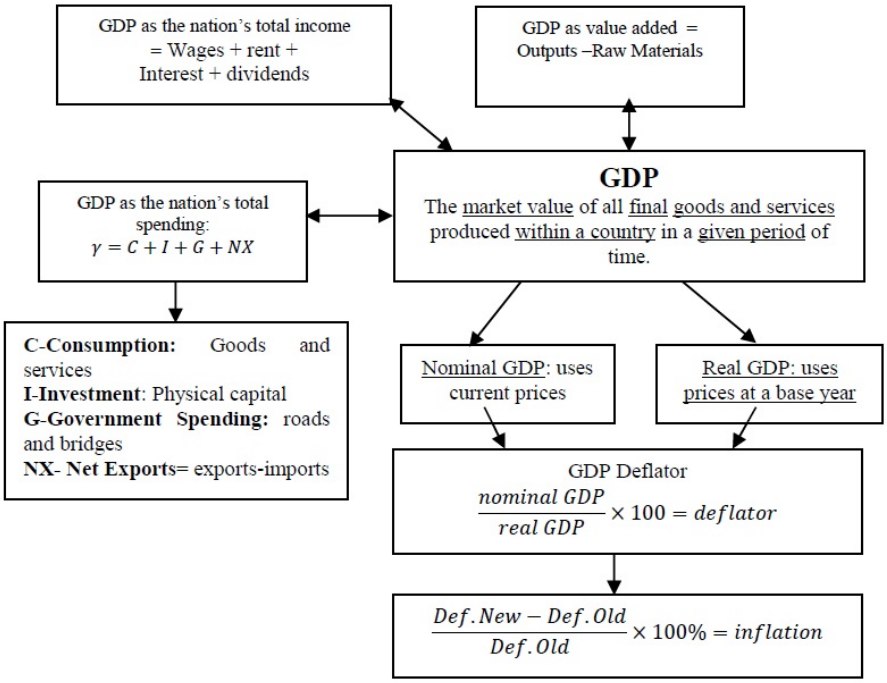


Figure 2. GDP: Conceptual Chapter Map



Figures 3 and 4 illustrate the differences between a conceptual chapter map and a traditional concept map for the concept of supply curve. Figure 3 shows a traditional concept map, adopted from Marangos and Alley (2007) and Figure 4 shows a Conceptual Chapter Map which I have used while teaching a Principles of Microeconomics course during the Spring 2013 semester at the University of Central Oklahoma.

Figure 3. Supply Curve: Traditional Concept Map

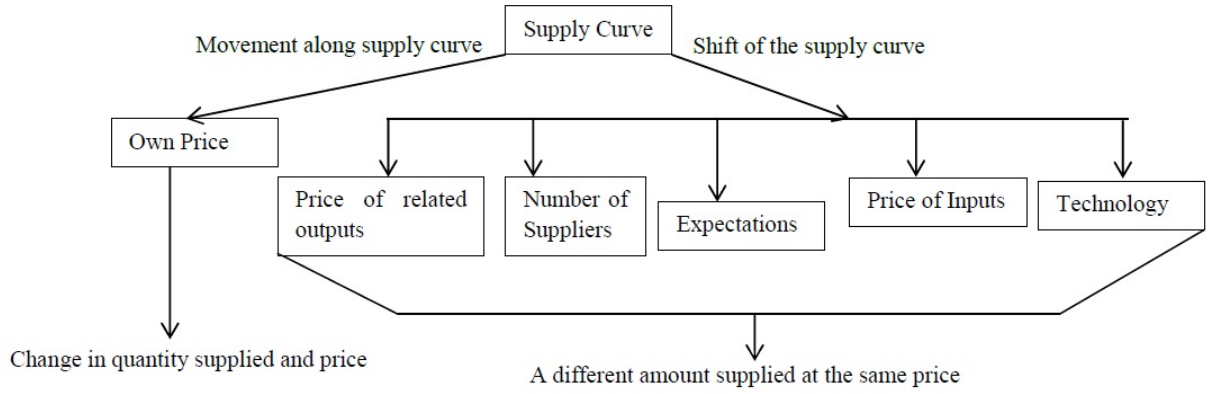


Figure 4. Supply Curve: Conceptual Chapter Map

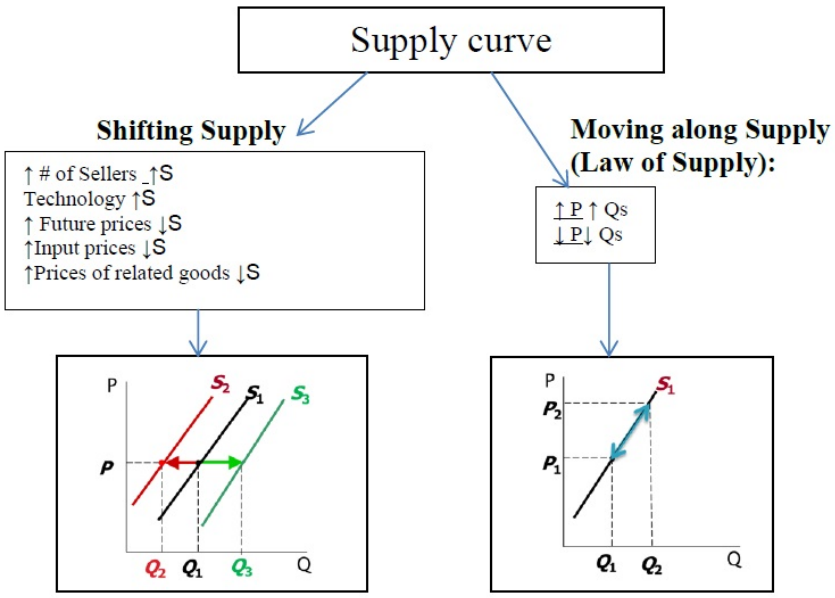


Figure 3 and figure 4 illustrate the same links among concepts related to supply curve. Both maps show that there is a difference between shifting supply curve and a movement along supply curve, and both indicate what causes these changes. In addition, the Conceptual Chapter Map also illustrates these changes using graphs, while the traditional concept map does not.

PRACTICAL STRATEGIES FOR CREATING AND DELIVERING CONCEPTUAL CHAPTER MAPS

While working with students in introductory economics courses, I have noticed certain patterns in students' inability to differentiate between concepts. While a visual learning tool such as a traditional concept map allows students to see the differences between concepts, it does not provide the context necessary to explain those differences. For example, a traditional concept map may illustrate that Price Elasticity of Demand is different from Income Elasticity, but it does not explain the underlying differences between those two concepts. Adding another layer of explanation provides students with contextual information they can use as a building block to develop further understanding of the concepts being introduced. The Conceptual Chapter Map was developed as a means to improve upon a traditional concept map by adding the missing contextual information to existing visual cues.

Starting in the fall of 2011 I have been using Conceptual Chapter Maps in Principles of Microeconomics and Principles of Macroeconomics courses. I would use the Conceptual Chapter Maps to conduct a review before in-class assignments, or I would ask students to create maps either in groups during the class or on their own after the class.

Research evaluating the effectiveness of teacher- vs. student-created maps has shown mixed results. Some studies show that students receive greater benefits by creating their own concept maps (Novak 1990). Others find that concept maps benefit students regardless of

whether the map is created by a student or instructor (Armbruster, Anderson, & Meyer 1991; Doyle 1999; Ellis 2001).

Instructors may also chose to assign students with the creation of concept maps in order to actively engage students in the learning process. When students create concept maps in groups they participate in collaborative learning, which has shown to have a positive effect on learning outcomes (Budd 2004; Baitz 2009). Student-created concept maps can also be used as a tool to evaluate students' understanding of a topic or a concept (Ruiz-Primo, Schultz, Li, & Shavelson 2001; Markham, Mintzes, & Jones 1994).

While the benefits are notable, there may also be downsides to student-created maps. First, most students are unfamiliar with the idea of concept mapping. Therefore, instructors need to spend time training students on how to develop a concept map. In fact, studies have shown that explicit training in the creation of concept maps is an important factor in its effective use (Lenz, Alley, & Schumaker 1987; McCagg 1991). This can be a time-consuming exercise and given the limited nature of class time, may not always be practical.

Second, students can resist learning how to create and use concept maps. From my personal experience, unless students receive a grade for creating a concept map, many may completely disregard the assignment. In their 2007 study, John Marangos and Sean Alley assigned students to create concept maps in a Principles of Microeconomics course. Study results revealed that students did not highly value the concept maps, and the majority of students completed the assignment only for the grade received. Most students stated they would likely not use this tool in the future.

Finally, students may misunderstand concepts or links among the concepts since the information presented in introductory economics classes is new for most. As a result, this may create an undesirable situation in which students visualize connections among concepts that may be wrong. Instructors can follow students' work during the assignment and try to correct these misperceptions in class, but given the limited nature of class time, this may not be feasible. Instructors can also collect student-created maps in order to make necessary corrections, but there is no guarantee that students will learn from these corrections.

Instead of making students responsible for creating a Conceptual Chapter Map, instructors can use a mixed approach that involves creating a map in class with the help of students. In order to avoid misunderstanding concepts and their connections, the instructor can create the "skeleton" of the map, and students can help fill it in with the definitions and explanations of concepts. For example, an instructor will show the difference between a shift in the supply curve and the movement along the curve, and students will have to answer which factors are causing which of the changes. This exercise creates a dialog between students and the instructor, and involves students in the process of map creation.

In my classes I have noticed that students trust instructor-created maps much more than they trust the maps created either on their own or in groups. When the maps are presented to students as a review before in-class assignments, students actually use them while working on the assignment. The ability of students to use a Conceptual Chapter Map while studying or working on an in-class assignment is one of its main benefits compared to traditional concept maps or traditional lecture notes.

From observing classroom behavior, instructors may notice that students are not relying on their notes when answering questions or when assigned to work on practice problems. Notes usually contain too much information and are thus difficult to use when trying to find answers. On the other hand, traditional concept maps may not have all of the explanations to make it

possible to answer the questions. Conceptual Chapter Maps are more practical since they combine the visual construct of a traditional concept map with the concept explanations found in the lecture notes.

Regardless of the method used, there will always be students who chose not to look for the answers among the tools provided, but instead rely upon their own, often incorrect, understanding of the subject. When answering questions from these students, the instructor can rely on the Conceptual Chapter Map to show students how the answer can be found or confirmed.

Since Conceptual Chapter Maps combine elements of both traditional concept maps and traditional lectures, instructors must be wary of providing information that is superfluous to understanding the basic concepts. Research shows that people learn better when extraneous material is excluded rather than included (Mayer 2009). Traditional lecture notes often have multiple definitions, examples and explanations, but not all of these are critical to understanding the concepts being taught. This creates an overload of information for the student, who is unable to discern what is and isn't most important. A Conceptual Chapter Map should be constructed in such a way that they are easily readable and contain only information essential for students to learn.

In this regard, the creation of a Conceptual Chapter Map is a useful exercise for the instructor. It helps to reorganize the material in a more meaningful way, and can improve lecture delivery by allowing the instructors to evaluate which points are most important and should be emphasized. Additionally, when compared to a traditional concept map, a Conceptual Chapter Map forces the instructor to rethink and reevaluate how to explain concepts during the lecture.

STUDENT ATTITUDES TOWARDS THE CONCEPTUAL CHAPTER MAPS

During the Spring 2013 semester, students in Principles of Microeconomics classes at the University of Central Oklahoma were introduced to Conceptual Chapter Maps. At the beginning of the class in which students would work on the in-class assignments, I would draw a Conceptual Chapter Map on the board and would encourage students to refer to the map while working on the assignment. My role was to ensure that the most important concepts were included in the map, and that the map illustrated the correct links among those concepts. Students also participated in map creation by finding concept definitions, explanations, and illustrations. I helped rewrite these explanations in a shortened form. For example, students would say that Supply increases when more sellers enter the market and I would rewrite this statement as "↑# sellers →↑S." Rewriting concept explanations in a simplified form assured that the maps would not be overcrowded with information. If during the in-class assignment students had any questions, I would answer them by referring directly to the conceptual chapter map.

At the end of the semester, students were asked to complete a survey revealing their attitudes toward the conceptual chapter maps. The surveys were administered during the last class of the semester and submitted by students anonymously. Of a total 65 students comprising both Principles of Microeconomics courses, 51 students completed the survey. Of those, 46% were freshman, 40% were sophomores, and 14% were either juniors or seniors. The average GPA indicated by students was 3.02, and the average grade the students expected to receive in class was 2.92.

Most of the survey questions were true/false questions. Since students had access to both traditional lecture notes and conceptual chapter maps, they were asked to specify which tool they

preferred to use while studying. Students were also asked to examine the concept maps they received throughout the semester and specify which chapter maps they found to be most useful and least useful. The survey questions are presented in the Appendix C. The results of the Survey are presented in Table 1.

Table 1
Students' Attitudes Towards Conceptual Chapter Maps

	% answered True
Conceptual Chapter Maps were helpful	96
Conceptual Chapter Maps were easy to use	90
Conceptual Chapter Maps helped me learn	94
I have used Conceptual Chapter Maps while working on Homework	70
I have used Conceptual Chapter Maps while working on in-class assignment	72
I have used Conceptual Chapter Maps while studying for midterm	94
I have used Conceptual Chapter Maps while studying for a quiz	70
I will use Conceptual Chapter Maps to study for the Final	92

Overall, the survey results indicated that students had a positive regard for the conceptual chapter maps. Questions 1-3 investigated whether the conceptual chapter maps were considered to be a useful studying tool by students. Survey results showed that 94% of students agreed that the concept maps helped them learn. Ninety-six percent of students indicated that conceptual chapter maps were helpful, and 90% of students indicated that the conceptual chapter map was easy to use.

Of the results reported, 71% of students stated they used concept maps while working on in-class assignments, during a homework assignment, or while studying for a quiz. The vast majority of students (94%) stated that they used the conceptual chapter map while studying for the midterm, and 92 % of students stated that they would use the map to study for the final exam. Marangos and Alley (2007) asked their students similar questions, but the results were not as favorable toward the traditional concept maps used in their classroom. They found that on average, 43% of students used concept maps to study for midterm and 62% of students indicated they would use concept maps in their final exam preparation.

The University of Central Oklahoma students surveyed were also asked whether they preferred to use conceptual chapter maps or lecture notes while studying. Of all students reporting, 45% stated that they prefer conceptual chapter maps to the lecture notes, 20% prefer traditional lecture notes, and 35% preferred to use both. This result indicates that conceptual chapter maps are not a perfect substitute for the lecture notes, but rather an additional tool students can refer to while studying.

Table 2
The Most and the Least Useful Conceptual Chapter Maps

	Most Useful (# of mentions)	Least useful (# of mentions)
Production Possibility Frontier	0	10
Supply and Demand	25	5
Elasticity	25	6
Government controls	24	3
Welfare economics and its application	32	4
Costs	23	2
Market structure	17	0

The results from table 2 are somewhat surprising but also informative. No students found the conceptual chapter map for Production Possibility Frontier to be most useful, but 20% found it to be the least useful. This may indicate that in the beginning of the semester, when students were first being introduced to conceptual chapter maps, they did not fully understand or appreciate the benefits of this tool. The other possible explanation is that the concepts related to PPF were not as widely used throughout the semester as Supply and Demand concepts. Finally, students may not have had difficulties understanding these concepts, thus explaining why they regarded this particular Conceptual Chapter map as the least useful. Other maps were somewhat similar in their usefulness, with most students indicating their usefulness rather than the opposite. While it is noticeable that different students regard different conceptual chapter maps as being more useful than others, the overall results demonstrate that if the time allows, it's better to provide students with a chapter map regardless of its perceived usefulness than not.

Students were also asked to provide comments regarding usefulness of concept maps. Appendix B provides the list of all comments received from students. Interestingly enough, even though the benefits of the conceptual chapter maps were never discussed in the classroom, many students in their comments pointed out the benefits of conceptual chapter maps similar to those as described in the paper. To summarize, the following benefits of the conceptual chapter maps were mentioned the most by the students:

Summarized main idea of each chapter
Showed links among concepts
Good for visual learners
Good study tool, easy and convenient
Good way to review material
Provided explanations, equations, and graphs
Concise summary of each chapter, one page instead of several pages of notes

There are several questions that require further investigation. First, students in my classes were only comparing conceptual chapter maps with traditional lecture notes. In the future, it would be interesting to investigate student attitudes towards conceptual chapter maps in relation to traditional concept maps. Also, the existing literature provides extensive reports on the effectiveness of traditional concept maps, but further investigation is needed to evaluate the effectiveness of the conceptual chapter maps in student learning outcomes.

CONCLUSIONS AND DISCUSSION

The goal of this paper was to introduce a practical teaching and learning tool which combines both the elements of traditional concept maps and lecture notes by illustrating the links among concepts and providing explanations of those concepts. Pulling from the idea of traditional concept maps, Conceptual Chapter Maps allow students to learn about conceptual relationships while concurrently utilizing definitions and explanations from lecture notes used while working on practical assignments or studying for exams.

Conceptual Chapter Maps can be delivered to students in the form of a topic review prior to a practical assignment. With conceptual Chapter Maps, students receive a condensed version of lecture notes, which helps them to review covered material and also allows them to organize and structure newly-obtained knowledge in a simple, visual way. Additionally, the Conceptual Chapter Map contains all the covered material on a single page so students do not need to flip through several pages of notes in order to find the necessary information to answer questions.

Creating Conceptual Chapter Maps is also a useful exercise for instructors, as it allows them to reorganize the material, evaluate which points are most important and should be emphasized, thereby improving upon lecture delivery.

The students studying Principles of Microeconomics at the University of Central Oklahoma during Spring 2013 semester responded extremely positively to the use of conceptual chapter maps as a learning tool. Perhaps visual learners benefited the most from using conceptual chapter maps, as students specifically noted that these maps helped them visualize learned material. In addition, most of the students stated they used conceptual chapter maps while studying for both midterm and final exams. Even those who said they preferred to use lecture notes for studying would still refer to conceptual chapter maps to clarify information. Based on students' comments, it can be stated that most students believe the conceptual chapter maps helped them learn. Whether conceptual chapter maps actually improve learning outcomes is yet to be determined.

APPENDIX A. QUESTIONER

This questioner is designed to identify the benefits of conceptual chapter maps in the class you are taking. To participate in this study, please fill out the questionnaire and return in to survey administrator. Your involvement is completely voluntary and your decision to participate or not participate will not influence your grade in this class. All information gathered for this study will be anonymous. This means that there is no identifying information linking you to the survey that you fill out.

GPA
Year in college
Time spent studying per week for this class
Expected grade in this class

Concept maps were helpful in general		Yes	No
Concept maps were easy to use		Yes	No
I used concept maps while			
Doing homework		Yes	No
Studying for a quiz		Yes	No
Studying for a midterm		Yes	No
Working on in-class assignment		Yes	No
Concept maps helped me learn		Yes	No
While studying, I prefer to use:	Maps	Lecture notes	Both

Please review concept maps and answer the following questions:

List numbers of chapters for which concept maps were the most useful

List number of chapters for which concept maps were the least useful

Please comment on usefulness of chapter maps in Principles of Microeconomics class:

APPENDIX B.
STUDENTS' COMMENTS REGARDING CONCEPTUAL CHAPTER MAPS

It was most helpful to my understanding of differences among graphs
The concept maps were very useful because they were the main idea of lecture so you hit the main points and put them in one paper
I liked using them, especially when preparing for midterm/final
They were helpful for review before tests and to make the note cards
I think they help you overall since they are easy to use and convenient
Made everything so much more clear. Making them was a way to recap everything and study!
The chapter maps helped me understand the material better, as well as helped me study for the midterm
I feel like if I just look at the maps I get confused. They were a great asset only after I had understood the material. For me they were a little overwhelming. They were still helpful.
The chapter maps have useful data + information but it is hard to interpret some in user friendly terms.
Lecture notes along with maps were all useful
Very useful, clear summary of a chapter
After learning and reading the material I used maps to review. Very useful to review before a quiz or a test
Maps took all the lecture information and condensed it into one map that helped show connections
They were clear and made studying easier
Breaks things down, makes them easier to understand
They were useful because I am very visual learner and the pictures made it easier for me to understand
They were all very useful and summarized the material into quick study tools
I feel like when I would use them during HW or studying, when I took the test I could picture the maps and better remember the material
They were useful, I just prefer lecture notes. But I would refer to chapter maps when get confused in notes.
Some were helpful but others were jumbled and didn't help that much. The ones that were helpful I really liked
They are very useful for studying for the tests in this class
Very useful and easy to comprehend. Helped when doing in-class assignments and homework
Extremely useful, especially having the equations needed and graphs, the directions.
They shortened the time of reviewing
Help to summarize the chapters
They helped me studying and figure out the chapters
They were a great help!
Good for visual learners
I wish we had more, and had blank ones on the test. My handwriting is bad, so having them posted online helped me to see what I failed to draw. Continue using the maps.
I like using lecture notes since they are more in depth and have graphs. Many times I used the maps but still had to use notes.
Visualized the information
They did help but some are more confusing than helpful
Maps come handy as quick references, in particular when solving problems
They gave overall vision of what to learn
At times I confused myself with a map but all in all I enjoyed having them. Would recommend hanging them out sooner
I am a visual learner so the maps really helped me to learn the material
The concept map helped me to understand the material a lot more than reading through my notes
It made it easier for me than when I used the lecture notes, clear and useful
I used the maps while studying but I feel like chapter 4 wasn't very helpful

I enjoyed the maps. They were concise summaries of each chapter and made studying easier and more efficient

All of the maps were useful. It makes it easier to read when the notes are visually broken down in graphs

All of the maps were very useful to me

It is very convenient and the basics of a chapter can be found on one piece of paper

REFERENCES

Armbruster, B. B., Anderson, T. H., & Meyer, J. L. (1991). Improving content-area reading using instructional graphics. *Reading Research Quarterly, 2*6(4), 393-416.

Ausubel, D.P. (1960). The use of advance organizers in the learning and retention of meaningful verbal material. *Journal of Educational Psychology*, 51, 267-272.

Baitz, I. (2009). Concept mapping in the online learning environment: a proven learning tool is transformed in a new environment. *International Journal of Learning*, 16 (8): 285-291.

Budd, J. W. (2004). Mind maps as classroom exercises. *Journal of Economic Education* 35 (1): 35.

Chiou, Chei-Chang. (2008) The effect of concept mapping on students' learning achievements and interests. *Innovations in Education and Teaching International*, Vol. 45, No. 4: 375–387

Doyle, C. S. (1999). The use of graphic organizers to improve comprehension of learning disabled students in social studies. Union, NJ: M. A. Research Project, Kean University. (ERIC Document Reproduction Service No. ED427313).

Ellis, E. S. (2001). *Makes Sense Strategies: Connecting Teaching Learning and Assessment Computer software.* Tuscaloosa, AL: Masterminds Publishing.

Horton, P.B., McConney A.A., Gallo M., Woods A.L., Senn G.J., Hamelin D. (1993). An investigation of the effectiveness of concept mapping as an instructional tool. *Science Education* 77 (1): 95-111.

Jacobs-Lawson, J.M., & Hershey, D.A. (2002). Concept maps as an assessment tool in psychology courses. *Methods & Techniques, 29*(1), 25–29.

Johnston, Carol, Concept Mapping For Students and the Faculty of Business and Economics Teaching and Learning Unit, Faculty of Business and Economics, University of Melbourne.

Lenz, B. K., Alley, G. R., & Schumaker, J. B. (1987). Activating the inactive learner: Advance organizers in the secondary content classroom. *Learning Disability Quarterly*,10(1) 53-67.

Marangos, J. (2003). The effectiveness of concept maps in introductory microeconomics. *Economic Papers* 22 (4): 74-82.

Marangos, J. and Alley, S. (2007), Effectiveness of concept maps in economics: evidence from Australia and USA. *Learning and Individual Differences* 17 (2): 193-199.

Markham, K.M., Mintzes, J.J., & Jones, M.G. (1994). The concept map as a research and evaluation tool: Further evidence of validity. *Journal of Research in Science Teaching, 31*(1), 91- 101.

Mayer, R. E. (2009), *Multimedia learning* (2nd ed). New York: Cambridge University Press.

McCagg, E.C., & Dansereau, D.F. (1991). A convergent strategy for examining knowledge mapping as a learning strategy. *Journal of Educational Research, 84*, 317–324.

Novak, J.D. (1990). Concept maps and vee diagrams: Two metacognitive tools for science and mathematics education. *Instructional Science, 19*, 29-52.

Novak, J.D. and Canas A.J. (2008). The theory underlying concept maps and how to construct and use them. Technical report IHMC CmapTools (2006-01; Rev 01-2008). Florida Institute for Human and Machine Cognition. http://cmap.ihmc.us/Publications/ResearchPapers/TheoryUnderlyingConceptMaps.pdf.

Preszler, R. (2004). Cooperative concept mapping improving performance in undergraduate biology. *Journal of College Science Teaching, 33*, 30-35.

Ruiz-Primo, M. A., Schultz, S. E., Li, M., & Shavelson, R. J. (2001). Comparison of the reliability and validity of scores from two concept-mapping techniques. *Journal of Research in Science Teaching, 38*(2), 260-278.

Willerman, M. and R.A. Macharg (1991). The Concept Map as an Advance Organizer. *Journal of Research in Science Teaching* 28 (8): 705-711.

QUALITIES UNIVERSITY STUDENTS SEEK IN A TEACHER

Shishu Zhang, University of the Incarnate Word
David Fike, University of the Incarnate Word
Gladys DeJesus, Chowan University

ABSTRACT

The unique contribution of the study is that it explores the perception of an ideal teacher from students of different backgrounds rather than evaluations of faculty performance. The paper finds that the two qualities that ranked highest at both schools were "knowledgeable" and "grades fairly." Another very interesting result is that faculty's use of technology was ranked to be the least important quality at both universities. Females rated five of the twenty-two faculty characteristics higher than did males. "At risk" students may need more support such as demonstrating a caring attitude, listening carefully to students, being sensitive to diversity, and engaging the students at their level.

INTRODUCTION

Wright (1997) documented that the most important factor that affects student learning is the teacher. Similarly, Kwan (1999) argued that more could be done to improve education by raising the effectiveness of teachers than by changing any other single factor.

Since faculty effectiveness is a concern, faculty and administrators are trying to upgrade faculty evaluation procedures to improve teachers' quality. Examining how to improve the effectiveness of the faculty evaluation process, Sporeen & Mortelmans (2006) and other scholars pointed out that apart from students' evaluation, other forms of evaluation should be used as complementary tools, such as (1) supervisor ratings, (2) self-ratings, and (3) peer ratings (Morgan, et al. 2003; Gursoy & Umbreit, 2005). Wright (2006) argues that current faculty evaluations are weak in providing teachers direction on how to improve their teaching or on why students respond as they do. Thus, the teaching evaluation should be improved so that it not only evaluates the faculty, but also provides directions as how to improve. In the same way, Steiner & Holley (2006) argued that teacher evaluation processes should not be an end-of-semester one-time evaluation, but instead should include, as a major component, a reliable and valid measure of a teacher's effects on student academic growth over time.

Thus, many researchers argue that we need to have diversified ways to obtain feedbacks from students about faculty effectiveness. This paper studies students' perceptions of the relative ranking of qualities of an ideal teacher in two different universities. It contributes to the debates about teacher qualities, teaching assessment and teacher effectiveness by providing the students' perception of an ideal teacher.

LITERATURE REVIEW

Teaching Evaluation for Different Groups

Teaching evaluations done by students are generally considered valid tools to measure faculty's performance and are widely used in universities.

One group of researchers has analyzed the relationship between the teacher evaluation ratings and the perception of the instructor as a strict or lenient grader. Crumbley (2001) claimed that students use teaching evaluations to punish instructors for being strict in grading, for giving a lot of quizzes and homework, and for asking questions that students cannot answer. Bacon and Novotny (2002) found that students who strive for achievement rank teachers who grade strictly higher than teachers who grade leniently. Marsh (2001) and Gursoy and Umbreit (2005) found that students appreciate learning and hence there would be a positive relation between good workload and teaching evaluation ratings.

Different student bodies or different disciplines can also affect the results of students' evaluations. Wright (1997) argued that homogeneity or heterogeneity of students' ability levels could affect teaching evaluation: teachers who teach classes that are more heterogeneous than homogeneous in ability levels are perceived to produce lesser effects on student learning and receive lower scores on teaching evaluations. Marsh & Dunkin (1992) found that faculty in science and natural science disciplines are frequently rated lower than are faculty in humanities. Similarly, Beran & Violato (2005) showed that courses in the social sciences receive higher ratings than courses in the natural sciences and that lab-type courses receive higher ratings than lectures or tutorials.

Other researchers have analyzed the relationship between instructors' personality and their student evaluations. Emery et al. (2003) argued that popularity and personality traits affected the results of teaching evaluations and turned teaching evaluations into a popularity contest. Others have found that students correlated teaching competence and ability with the instructor's personality and gave higher evaluations to instructors who were seen as "kind and caring" and who had a "very positive feeling towards the class and students" (Kim et al. 2000) and to instructors whom they perceived as supportive and enthusiastic about the class and subject matter (Bacon, & Novotny, 2002; Gursoy & Umbreit, 2005). Some authors have argued that student evaluations of faculty should not be used in decisions for tenure and promotion because charismatic and enthusiastic instructors were found to earn high ratings, even if they may have had low knowledge of the subject matter (Emery, et al, 2003).

Thus many studies have examined different influences on teaching evaluations. Research has shown that teaching evaluations are biased by various factors, such as the student body, type of courses, workload, grades, and instructor's personality (Morgan, et al. 2003; Simpson, & Siguaw, 2000; Sojka, et al. 2002, MacDermott, 2013).

Gender Effect on Students' Evaluation

Researchers have also found differences in faculty based on the gender of the instructor. Female instructors were perceived as warmer and more patient individuals but were expected to offer greater interpersonal support and were judged more strictly than male instructors in providing it (Sheila, 1982). Students reported that female professors gave them more time and personal attention than did male professors (Crawford & MacLeod, 1990). But Smith, Yoo, Farr, Salmon, & Miller (2007) reported evidence from a large study of more than 12,000 students that there was

no difference by faculty gender on student course evaluations. Similarly, Basow (2000) found that there was no gender difference in student perceptions of one's "best" professor. Smith (2007) found that for both male and female students, best professors exhibited both masculine traits and feminine traits (Smith, 2007) and further suggested that administrators should not assume one sex to provide better or worse instruction and that they should reward instructors on the basis of individual performance rather than according to the instructor's gender.

However, research suggests that although direct gender bias may not be observed in formal student evaluations of their instructors, female faculty members are nonetheless subject to culturally conditioned gender stereotypes (Feldman, 1993). As for the different responses from the female and male students, in a study comparing students from the United States, Jordan, and Chile, Alshare and Miller (2009) found that male and female American students differed in their perception of traits assessing teaching style and class management: female students prefers a more warm and care teaching style while male students prefers a sober, quiet learning environment.

Ideal Teacher

In contrast to teaching evaluation literature, there is very limited research on students' perception of an ideal university teacher. Initial research by Rubin (1981) on the "ideal professor," students identified five general categories of traits they rated highly. These five categories included items relevant to (a) knowledge, intellect, and ability, or expertise in the subject; (b) professionalism, or qualities that command respect; (c) ability to communicate; (d) openness (to students and their ideas); and (e) being nurturing and supportive.

More recently, Strage (2008) found that the most frequently-cited characteristics of an ideal professor include being knowledgeable, caring, concerned about students, and funny or entertaining. Helterbran (2008) used "ratemyprofessor.com" to study students' perceptions of the ideal professor. He presented these categories of characteristics as important: knowledge and presentation, and professional personal qualities (enthusiasm, approachability, caring).

Strage (2008) argued that students at Christian colleges or universities may hold different expectations of professors than their peers at public or private universities, given the centrality of the role of instructor in upholding and living out the mission of such institutions. He noted that despite the rapid growth of private Christian institutions, students and professors at religiously affiliated institutions have not conducted much research related to student expectations of the ideal professor.

Woods, Badzinski, Fritz & Yeates's (2012) research is one of the studies about students' perception of an ideal professor conducted in a Christian university. They administered a survey to 451 undergraduate students at a private liberal-arts Christian university, which revealed that their ideal professor places great emphasis on the integration of faith and learning, is flexible, maintains high academic standards, encourages students, and has an adaptive teaching style. Their findings also highlighted gender differences in student perception of the ideal professor. Female students ranked an adaptable teaching style, encouragement, and integration of faith and learning as more important than did male students.

STUDY PURPOSE

Similar to these studies, the purpose of this study is to determine the qualities of university faculty that economics students from two universities report as being important. The research questions are:

- *What are students' rankings (relative importance) of university faculty qualities from the two universities?*
- *Do students' ratings of faculty qualities differ by schools?*
- *Do students' ratings of faculty qualities differ based on students' GPA, gender, or parents' education?*

Students' expectations of faculty were compared among students at the two universities. The unique feature of this research is that it doesn't study teaching evaluations of specific courses; therefore, the results were unaffected by the possible biases such as students' expected grades, the course workloads, and the difficulty of the course. In addition, in order to get more robust results, the study compared the perceptions of ideal faculty in two different universities.

METHODS

Setting

This study was conducted at two universities in 2011-2013. The first university is located in the southwest US, which offers undergraduate and graduate degrees including the Ph.D. and professional degrees. This private, urban university is a Hispanic Serving Institution (HSI) with an enrollment of about 9000 students. The second institution is also a private university, but located in a small town in the southeast US. It awards the associate's, bachelor's and master's degrees in a variety of disciplines, and has an enrollment of 1200 students. Sixty-seven percent of the students are African American.

Study Design

This study entailed a cross-sectional survey of economics students each semester. The first university observed students for 4 semesters while the second university gathered the data for 2 semesters. An instrument (see Table 1) was developed by faculty at one university to determine the students' perspectives on the importance of faculty qualities. Students attending classroom economics courses (no online courses were included) were asked to rate each of the twenty-two faculty qualities listed on the instrument. The students completed the hardcopy instrument in class. The students ranked each quality on a Likert-type scale with 1 to 5 corresponding to "not important at all," "not very important," "neutral," "important," and "very important." Two institutions were included in the study to determine if the importance of faculty qualities were roughly consistent between economics students at the two institutions. Institutional Review Board approval at both universities was received before the study commenced.

Sample

The sample (see Table 2) was comprised of 458 economics students from the two institutions. Graduate students and upperclassmen comprised 55% of the sample. The majority of the students had parents who were not college graduates. Less than one-third of the students were from the second university, with 71% attending the first university. Males comprised 56% of the sample.

Instrument Reliability

Following development of the original instrument, it was peer-reviewed by faculty for content validity and pilot-tested with a group of students taking economics courses. Based upon the feedback, the instrument was revised. Table 1 lists the items in the revised instrument. Reliability of the revised instrument was assessed for the sample of 458 students. Cronbach's alpha for the 22-item instrument was 0.894, indicating appropriate internal consistency (George & Mallery, 2003).

Table 1. Teacher Qualities
The teacher:
1. Is knowledgeable about the subject.
2. Conveys knowledge effectively.
3. Is punctual.
4. Demonstrates confidence when teaching.
5. Is patient with students.
6. Demonstrates a caring attitude for students.
7. Listens to students' concerns.
8. Is dedicated to excellence in teaching.
9. Is sensitive to diversity.
10. Offers adequate resources for students to complete tasks.
11. Is willing to help students achieve goals.
12. Encourages students to achieve.
13. Shows passion for life.
14. Makes the classroom fun.
15. Relates classroom work with life examples.
16. Uses new technology to teach.
17. Presents materials in an organized way.
18. Grades fairly.
19. Sets attainable expectations for the students.
20. Is dedicated to the Mission of the university.
21. Engages students in the classroom.
22. Is intellectually stimulating.

Statistical Analyses

The data were analyzed using IBM SPSS Statistics 21 (IBM, Armonk, NY). Descriptive statistics were produced, including N's, percentages, means, and standard deviations. Correlation coefficients were generated, including Pearson's for continuous data and Spearman's for ordinal data. To compare means between two groups, independent sample t-tests were produced. Standardized effect sizes (Cohen's d) were calculated. Unless noted otherwise, the *a priori* level of significance was .05.

Table 2. Sample Characteristics for 458 Students[a], N (%) or Mean ± SD	
Student GPA	3.27 ± 0.55
Freshman	75 (16)
Sophomore	130 (28)
Junior	85 (19)
Senior	44 (10)
Graduate	117 (26)
Female	202 (44)
Male	255 (56)
Parent Graduated College, Yes	211 (46)
Parent Graduated College, No	239 (52)
Students of the first university	324 (71)
Students of the second university	134 (29)
[a] There are missing data elements for < 2% of the student sample.	

RESULTS

Table 3 reports the ranking of teachers' qualities from both universities. The two highest-ranked qualities as rated by the students from the first university are "knowledge" and "grade fairly." The top two qualities of faculty rated by students from the second university are the same as the first university, but the order of the ranking is different. The top five qualities are the same except that the first school ranked "confidence" in the top five qualities while the second school ranked "encourages students" there instead. Besides this difference, the orders of the ranking of the top five qualities for the two schools are different. Some of the lowest rated faculty qualities include "dedicated to mission," "passion for life," "makes class fun," "uses new technology," and "sensitive to diversity." A very interesting result is that "using new technology" is ranked as the lowest quality for students in both universities, which is consistent with our previous findings using data from Economics and Education students only at the first university.

Table 3. Rating of Faculty Qualities, Rank Ordered by Means			
First University		Second University	
Quality	Mean ± SD	Quality	Mean ± SD
Q1. Knowledgeable	4.84 ± .408	Q18. Grades Fairly	4.80 ± .484
Q18. Grades Fairly	4.80 ± .443	Q1. Knowledgeable	4.72 ± .667
Q2. Conveys Knowledge	4.77 ± .467	Q11. Willing To Help	4.72 ± .597
Q11. Willing To Help	4.67 ± .556	Q12. Encourages Students	4.70 ± .696
Q4. Confident	4.60 ± .583	Q2. Conveys Knowledge	4.69 ± .652
Q12. Encourages Students	4.58 ± .669	Q7. Listens	4.67 ± .648
Q19. Sets Attainable Goals	4.56 ± .599	Q19. Sets Attainable Goals	4.67 ± .562
Q8. Dedicated To Excellence in Teaching	4.54 ± .688	Q8. Dedicated To Excellence in Teaching	4.66 ± .673
Q17. Organized	4.53 ± .620	Q17. Organized	4.60 ± .651
Q5. Patient	4.51 ± .670	Q4. Confident	4.59 ± .696
Q10. Offers Resources	4.51 ± .661	Q6. Caring	4.59 ± .628
Q7. Listens	4.48 ± .706	Q22. Intellectually Stimulating	4.56 ± .671
Q22. Intellectually Stimulating	4.44 ± .714	Q10. Offers Resources	4.47 ± .765

Table 3. Rating of Faculty Qualities, Rank Ordered by Means			
First University		Second University	
Quality	Mean ± SD	Quality	Mean ± SD
Q6. Caring	4.43 ± .737	Q21. Engaging	4.47 ± .661
Q15. Uses Real Life Examples	4.33 ± .854	Q5. Patient	4.46 ± .783
Q21. Engaging	4.27 ± .823	Q20. Dedicated To Mission	4.39 ± .860
Q14. Makes Class Fun	4.24 ± .853	Q15. Uses Real Life Examples	4.36 ± .860
Q3. Punctual	4.23 ± .784	Q3. Punctual	4.34 ± .824
Q9. Sensitive To Diversity	4.14 ± 1.01	Q9. Sensitive To Diversity	4.25 ± .987
Q13. Passion For Life	4.03 ± .976	Q13. Passion For Life	4.17± .955
Q20. Dedicated To Mission	4.01 ± .994	Q14. Makes Class Fun	4.02 ±1.055
Q16. Uses New Technology	3.73 ± 1.03	Q16. Uses New Technology	3.97±1.058

Table 4 reports the difference of ranking between the two universities for each item. We found that the two schools' rankings are significantly different for seven out of the twenty-two items. "Knowledgeable" is highly ranked characteristic for both schools; however, the first university's students ranked "knowledgeable" and "makes class fun" significantly higher than the second university, while the second group rated "caring," "listens," "uses new technology," "dedicated to mission," and "engaging" significantly higher.

Table 4. Rating of Faculty Qualities by School (Mean ± SD)				
Quality	First University	Second University	p-value	Cohen's d
Q1. Knowledgeable	4.84 ± .408	4.72 ± .667	.048*[u]	0.217
Q2. Conveys Knowledge	4.77 ± .467	4.69 ± .652	.195 [u]	0.141
Q3. Punctual	4.23 ± .784	4.34 ± .824	.150	-0.137
Q4. Confident	4.60 ± .583	4.59 ± .696	.862	0.016
Q5. Patient	4.51 ± .670	4.46 ± .783	.527 [u]	0.069
Q6. Caring	4.43 ± .737	4.59 ± .628	.015* [u]	-0.234
Q7. Listens	4.48 ± .706	4.67 ± .648	.007* [u]	-0.280
Q8. Dedicated To Excellence in Teaching	4.54 ± .688	4.66 ± .673	.094	-0.176
Q9. Sensitive To Diversity	4.14 ± 1.01	4.25 ± .987	.267	-0.110
Q10. Offers Resources	4.51 ± .661	4.47 ± .765	.603	0.056
Q11. Willing To Help	4.67 ± .556	4.72 ± .597	.367	-0.087
Q12. Encourages Students	4.58 ± .669	4.70 ± .696	.103[u]	-0.176
Q13. Passion For Life	4.03 ± .976	4.17 ± .955	.179	-0.145
Q14. Makes Class Fun	4.24 ± .853	4.02 ± 1.06	.035* [u]	0.229
Q15. Uses Real Life Examples	4.33 ± .854	4.36 ± .860	.699	-0.035
Q16. Uses New Technology	3.73 ± 1.03	3.79 ± 1.05	.026*	-0.058
Q17. Organized	4.53 ± .620	4.60 ± .651	.847	-0.110
Q18. Grades Fairly	4.80 ± .443	4.80 ± .484	.965	0
Q19. Sets Attainable Goals	4.56 ± .599	4.67 ± .562	.082	-0.189
Q20. Dedicated To Mission	4.01 ± .994	4.39 ± .860	<.001*	-0.409
Q21. Engaging	4.27 ± .823	4.47 ± .661	.007* [u]	-0.268
Q22. Intellectually Stimulating	4.44 ± .714	4.56 ± .671	.103	-0.173
* Statistically significant at 0.05 level. [U] T-test with unequal variances.				

Table 5 provides a comparison of ratings by students who have at least one parent with college degree with those whose parents do not have college degrees. For most of the faculty-quality ratings, the two groups of students do not differ. However, students who have a parent with college degree rated "dedicated to excellence in teaching" significantly lower than students whose parents do not have a college degree.

Table 5. Rating of Faculty Qualities by Students' Parents' Education (Mean ± SD)				
Quality	Not Col Grad	Col Grad	p-value	Cohen's d
Q1. Knowledgeable	4.79 ± .508	4.82 ± .497	.608	-0.060
Q2. Conveys Knowledge	4.71 ± .576	4.79 ± .475	.129 [u]	-0.15
Q3. Punctual	4.27 ± .801	4.27 ± .796	.975	0
Q4. Confident	4.57 ± .650	4.63 ± .575	.307	-0.098
Q5. Patient	4.46 ± .744	4.54 ± .657	.256	-0.114
Q6. Caring	4.49 ± .673	4.47 ± .740	.763	0.028
Q7. Listens	4.53 ± .697	4.55 ± .699	.783	-0.029
Q8. Dedicated To Excellence in Teaching	4.65 ± .676	4.50 ± .694	.027* [u]	0.219
Q9. Sensitive To Diversity	4.13 ± 1.039	4.22 ± .974	.378	-0.089
Q10. Offers Resources	4.50 ± .704	4.50 ± .680	.949	0
Q11. Willing To Help	4.70 ± .568	4.67 ± .564	.602	0.053
Q12. Encourages Students	4.63 ± .679	4.59 ± .686	.515	0.059
Q13. Passion For Life	4.09 ± .979	4.05 ± .970	.658	0.041
Q14. Makes Class Fun	4.16 ± .959	4.21 ± .881	.576	-0.054
Q15. Uses Real Life Examples	4.34 ± .830	4.34 ± .854	.961	0
Q16. Uses New Technology	3.82 ± 1.027	3.80 ± 1.046	.885	0.019
Q17. Organized	4.60 ± .634	4.58 ± .574	.811	0.033
Q18. Grades Fairly	4.78 ± .491	4.83 ± .413	.223 [u]	-0.110
Q19. Sets Attainable Goals	4.61 ± .569	4.59 ± .607	.696	0.034
Q20. Dedicated To Mission	4.11 ± 1.001	4.15 ± .942	.653	-0.041
Q21. Engaging	4.36 ± .740	4.29 ± .831	.360	0.089
Q22. Intellectually Stimulating	4.48 ± .694	4.48 ± .722	.992	0
* Statistically significant at 0.05 [U] T- test with unequal variances.				

Table 6 compares the teachers' quality ratings by student gender. Female and male student ratings are significantly different for five out of the twenty-two items. The female students rate these faculty qualities significantly higher than the male students: "conveys knowledge," "confident," "patient," "sensitive to diversity," and "organized."

Table 6. Rating of Faculty Qualities by Student Gender (Mean ± SD)				
Quality	Male	Female	p-value	Cohen's d
Q1. Knowledgeable	4.78 ± .491	4.83 ± .513	.369	-0.100
Q2. Conveys Knowledge	4.70 ± .545	4.81 ± .503	.026* [u]	-0.210
Q3. Punctual	4.20 ± .834	4.34 ± .738	.059	-0.178
Q4. Confident	4.54 ± .662	4.66 ± .551	.039* [u]	-0.197
Q5. Patient	4.43 ± .729	4.57 ± .667	.031* [u]	-0.200
Q6. Caring	4.43 ± .746	4.53 ± .655	.129 [u]	-0.142
Q7. Listens	4.49 ± .743	4.59 ± .627	.140 [u]	-0.145
Q8. Dedicated To Excellence in Teaching	4.56 ± .720	4.60 ± .641	.542	-0.059
Q9. Sensitive To Diversity	4.06 ± 1.074	4.31 ± .896	.007*[u]	-0.253

Table 6. Rating of Faculty Qualities by Student Gender (Mean ± SD)				
Quality	Male	Female	p-value	Cohen's d
Q10. Offers Resources	4.49 ± .664	4.51 ± .728	.762	-0.029
Q11. Willing To Help	4.65 ± .568	4.72 ± .567	.206	-0.123
Q12. Encourages Students	4.57 ± .740	4.68 ± .590	.075 [u]	-0.164
Q13. Passion For Life	4.02 ± 1.025	4.13 ± .900	.252	-0.114
Q14. Makes Class Fun	4.16 ± .919	4.21 ± .923	.562	-0.054
Q15. Uses Real Life Examples	4.34 ± .797	4.35 ± .891	.920	-0.012
Q16. Uses New Technology	3.81 ± 1.026	3.79 ± 1.054	.815	0.019
Q17. Organized	4.53 ± .620	4.67 ± .576	.015* [u]	-0.234
Q18. Grades Fairly	4.79 ± .464	4.82 ± .444	.421	-0.066
Q19. Sets Attainable Goals	4.60 ± .595	4.59 ± .585	.912	0.017
Q20. Dedicated To Mission	4.11 ± 1.007	4.13 ± .929	.834	-0.021
Q21. Engaging	4.37 ± .756	4.27 ± .818	.222	0.127
Q22. Intellectually Stimulating	4.49 ± .690	4.45 ± .720	.537	0.057
[U] T-test with unequal variances.				
* Statistically significant at 0.05 level.				

Table 7 summarizes the correlations between teacher-quality ratings and students' GPA. Students' GPA is significantly correlated with eleven of twenty-two faculty-quality ratings. The higher the students' GPA, the higher they rated faculty qualities of "knowledgeable" and "conveys knowledge." However, the higher the students' GPA, the lower the students rate the faculty qualities of "punctual," "caring," "listens," "sensitive to diversity," "passion for life," "makes class fun," "uses new technology," "dedicated to mission," and "engaging."

Table 7. Correlation (r) of Rating of Qualities with Students' GPA

Table 7. Correlation (r) of Rating of Qualities with Students' GPA	
Quality	GPA
Q1. Knowledgeable	.179*
Q2. Conveys Knowledge	.197*
Q3. Punctual	-.100*
Q4. Confident	.002
Q5. Patient	-.025
Q6. Caring	-.119*
Q7. Listens	-.146*
Q8. Dedicated To Excellence in Teaching	-.034
Q9. Sensitive To Diversity	-.113*
Q10. Offers Resources	-.054
Q11. Willing To Help	-.092
Q12. Encourages Students	-.125
Q13. Passion For Life	-.198*
Q14. Makes Class Fun	-.130*
Q15. Uses Real Life Examples	-.065
Q16. Uses New Technology	-.174*
Q17. Organized	-.038
Q18. Grades Fairly	.088
Q19. Sets Attainable Goals	-.074
Q20. Dedicated To Mission	-.208*
Q21. Engaging	-.110*
Q22. Intellectually Stimulating	.014
* Statistically significant at 0.05 level	

DISCUSSION

This research was conducted at two universities. As shown in Table 3, the two qualities that ranked highest at both schools were “knowledgeable” and “grades fairly.” This particular finding may provide a strong incentive for university professors to continually improve their knowledge of subject and to be conscious of being fair in grading their students.

It is always a challenge to treat students equally, assigning fair grades to each one. “Grades fairly” may be interpreted differently by different students: some may consider “grades fairly” as “grades easily” while others may equate “grades fairly” to “grading consistently.” This study analyzes students’ perceptions of “fair” grading without differentiating among each student’s interpretation of the term. Regardless of students’ interpretations, this quality is one of the top two faculty qualities students considered most important.

Another very interesting result is that using technology was ranked as the least important faculty quality for students in both universities. This result was consistent with previous research about teaching qualities (Zhang, Fike & Fike, 2013). The result may suggest that our attempts at using new technology are seen as ineffective and that we fail to address the core task of teaching, which is to help students obtain and apply knowledge. In addition, students in both schools placed low importance on “makes class fun.” This result may serve as evidence against treating students as customers by entertaining them. The students in both schools also placed low importance on faculty qualities of “punctual,” “sensitive to diversity,” and “passion for life.”

Table 4 compares the ratings of teachers’ qualities at these two universities. The first, predominantly Hispanic, rated the qualities “knowledgeable” and “makes class fun” significantly higher than the second school did, while that school, predominantly African American, rated “caring,” “listens,” “engaging,” “uses new technology,” and “dedicated to mission” significantly higher instead. This result might indicate that racially different student bodies have different expectations for teachers. An ‘excellent’ teacher at one school might only receive an average student evaluation in another school, not because his or her teaching is not as good as before, but rather because of different expectations from racially different student bodies. Professors may need and should be given time to adjust their teaching style in order to adjust to these differences.

Table 5 demonstrates that students with college-degreed parents rated "dedicated to excellence of teaching" lower than students whose parents are without college degrees. This result may indicate that first generation students actually put more emphasis on teaching excellence than second- or third-generation college students. As presented in Table 6, females rated five of the twenty-two questions higher than males, suggesting that faculty might want to choose to emphasize different qualities depending on the gender ratio of the student body. Table 7 shows that the higher the current GPA, the higher the students’ ratings of “knowledgeable” and “conveys knowledge,” and the lower the ratings on other qualities. This result demonstrates that lower GPA students regardless of the institution where enrolled may need more support such as “caring”, “listens”, “sensitive to diversity”, and “engaging”

IMPLICATIONS FOR FUTURE PRACTICE AND RESEARCH

Students at both universities rated “knowledgeable,” and “grading fairly” as the top two qualities among the twenty-two faculty qualities. This result indicates that it is important for faculty to learn and generate new knowledge, and to use rigorous grading rubrics to ensure that

students can see that grading is fair. Consistent with the results of our former paper (Zhang, Fike & Fike, 2013), students at both universities rated "uses new technology" the lowest of the twenty-two qualities. From a practical perspective, this research suggests that the faculty who overemphasize the importance of using technology in the classroom might be better served by emphasizing other qualities. Future research could explore why the students rate using technology the lowest teacher quality.

The paper discovers that students' ratings of faculty qualities differ by universities. An implication for future practice is that if faculty members transfer from one university to another, they might obtain different teaching evaluations due to different students' expectations. Therefore, it is important to adjust their teaching styles and methods.

For faculty with students with higher GPAs, special attention should be paid to being knowledgeable and being able to effectively convey that knowledge; students with lower GPA may benefit if the faculty demonstrates other qualities such as "caring," "listens," "sensitive to diversity."

Female students have higher ratings for faculty in many categories than male students. This implies that for faculty with a larger percentage of female students, if they develop qualities such as "confident," "patient," "sensitive to diversity," and "organized," they might get better teaching evaluations. At the same time, the students may have a better learning experience/ outcome since they feel more comfortable in the classroom.

Students with college-degreed parents don't differ from other students in rating twenty-one of the twenty-two qualities. This suggests that the students' ratings of faculty qualities are not influenced by their parents' education level. An implication for practice is that the faculty may not need to treat students differently based on their parents' education level.

Additional research could be conducted to determine the reasons why students ranked certain faculty qualities high or low. For example, the students who ranked "uses new technology" low may think certain technology is not effective in improving learning outcomes. However, students may find some technology useful in improving teaching and learning. Future research could clarify which types of technology students consider beneficial.

This study compares students' ratings of faculty in two universities with a high percentage of students from traditional underserved populations. Future research could be conducted by comparing students' ratings of faculty from different disciplines across multiple universities, which could provide more generalizable findings.

LIMITATIONS

This study was conducted at two private universities that have a large percentage of students from traditionally-underserved populations. The sample for this study consisted solely of economics students. This cross-sectional, observational study did not employ a causal design.

Terms used in the instrument were not explicitly defined, so students' interpretations of the terms may vary. Though this was a potential limitation of the instrument, use of the instrument in this study achieved acceptable reliability as demonstrated by Cronbach's alpha = 0.894.

Student characteristics other than the demographic variables included in this study may have a bearing on ratings of faculty qualities. Further studies identifying other student characteristics that are associated with ratings of faculty qualities are needed. Additionally, research is needed to explore the reasons why students rate some faculty qualities as more important than others.

CONCLUSION

This study of 458 students in Economics disciplines at two universities resulted in very interesting and significant findings. Students at both schools ranked "knowledgeable" and "grades fairly" as top two qualities for faculty. Meanwhile, students in both universities ranked "uses new technology" as the least valued of the twenty-two faculty qualities assessed, which reinforced the results of our previous research.

We also found that the students' rankings of certain faculty qualities at the two universities are different; this may suggest that faculty members at different universities need to develop certain qualities or tailor their teaching style based on their students' values. Furthermore, the instrument used in the study could be a complement to students' teaching evaluations to help faculty develop qualities that students value most. For example, according to the result of this study, faculty may need to shift their focus from using technology and making class fun to conveying knowledge and grading fairly. Faculty who teach students with higher GPA need to pay special attention to becoming knowledgeable and being able to convey their knowledge. Faculty of students with lower GPA need to develop other qualities such as being caring and flexible in addition to being knowledgeable and being able to convey their knowledge. Members of faculty teaching female students need to develop other qualities besides being knowledgeable since female students rated other qualities of faculty more highly than male students. Knowing the expectations of students is an advantage to faculty who are pursuing effective teaching.

REFERENCES

Alshare, K. A., & Miller, D. S. (2009). Student perceptions of the importance of instructor traits: A cross-cultural study. *Academy of Educational Leadership Journal*, *13*, 107–134.

Bacon, D. R., & Novotny, J. (2002). Exploring achievement striving as a moderator of the grading leniency effect. *Journal of Marketing Education,* 24(4), 4-14.

Basow, S. A. (2000). Best and worst professors: Gender patterns in students' choices. *Sex Roles*, *43*, 407–417.

Bennett, S. K. (1982). Student perceptions of and expectations for male and female instructors: Evidence relating to the question of gender bias in teaching evaluation. *Journal of Educational Psychology*, 74(2), 170-179.

Beran, T., & Violato, C. (2005). Ratings of university teacher instruction: How much do student and course characteristics really matter? *Assessment and Evaluation in Higher Education Journal*, 30 (6): 593–601.

Crawford,M., &McLeod, M. (1990). Gender in the college classroom: An assessment of the "chilly climate" for women. *Sex Roles*, *23*, 101–125.

Crumbley, L., Henry, B. K., & Kratchman, S. H. (2001). Students' perceptions of the evaluation of college teaching. *Quality assurance in Education,* 9(4), 197-207.

Emery, C. R., Kramer, T. R., & Tian, R. G. (2003). Return to academic standards: a critique of student evaluations of teaching effectiveness. *Quality Assurance in Education,* 11(1), 37-46.

Feldman, K. A. (1993). College students' views of male and female college teachers: Part II Evidence from students' evaluations of their classroom teachers. *Research in Higher Education, 34,* 151–191.

George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference*, 11.0 update (4th ed). Boston: Allyn & Bacon.

Gursoy, D. & Umbreit, W.T. (2005). Exploring students' evaluations of teaching effectiveness: What factors are

important? *Journal of Hospitality Tourism Research*, 29(1), 91-109.
Helterbran, V. R. (2008). The ideal professor: Student perceptions of effective instructor practices, attitudes, and skills. *Education, 129*, 125–138.
Kwan, K. P. (1999) How fair are student ratings in assessing the teaching performance of university teachers? *Assessment & Evaluation in Higher Education*, 24 (2), 181–195.
MacDermott, R.J. (2013). The Impact of Assessment Policy on Learning: Replacement Exams or Grade Dropping, Journal of Economic Education, 44 (4), 364-371.
Marsh, H.W. (2001). Distinguishing between good (useful) and bad workloads on students' evaluations of teaching. *American Educational Research Journal*, 38(1), 183-212.
Marsh, H. W., & Dunkin, M. J. (1992). Students' evaluations of university teaching: A multidimensional perspective. *Higher education: Handbook of theory and research.* New York: Agathon Press.
Morgan, D.A., Sneed, J., & Swinney, L. (2003). Are students evaluations a valid measure of teaching effectiveness: Perceptions of Accounting faculty members and administrators. *Management research News*, 26(7), 17-32.
Rubin, R. B. (1981). Ideal traits and terms of address for male and female college professors. *Journal of Personality and Social Psychology*, 41(5), 966–974.
Simpson, P. M., & Siguaw, J. A. (2000). Student evaluations of teaching: An exploratory study of the faculty response. *Journal of Marketing Education*, 22(3), 199-213.
Sojka, J., Gupta, A. K., & Deeter-Schmelz, D. (2002). Student and faculty perceptions of student evaluations of teaching: A study of similarities and differences. *College Teaching*, 50(2), 44-49.
Smith, S.W. (2007). Student perceptions of and expectations for male and female instructors: Evidence relating to the question of gender bias in teaching evaluation. *Women's Studies in Communication*, 30 (1), 64-77
Smith, S., Yoo, J., Farr, A., Salmon, C., & Miller, V. (2007). The influence of student sex and instructor sex on student ratings of instructors: results from a college of communication. *Women's Studies in Communication*, 30(1), 64-77.
Sporeen, P. & Mortelmans, D. (2006). Teacher professionalism and student evaluation of teaching: Will better teachers receive higher ratings and will better students give higher ratings? *Educational Studies*, 32(2), 201-214.
Steiner, S., & Holley, L. C. (2006). Evaluating teaching: Listening to students while acknowledging bias. Journal of Social Work Education, 42(2), 355-376.
Strage, A. (2008). Traditional and non-traditional college students' description of the "ideal" professor and the "ideal" course and perceived strengths and limitations. *College Student Journal*, *42*, 225–231.
Wright, P. (1997). Teacher and Classroom Context Effects on Student Achievement: Implication for Teacher Evaluation. *Journal of Personnel Evaluation in Education* 11: 57-67.
Wright, R. E. (2006). Student evaluation of faculty: concerns raised in the literature, and possible solutions. College *Student Journal*, 40(2), 417-422.
Zhang, S., Fike, D.S., & Fike, R. (2013). Qualities of an Ideal Teacher: The Student Perspective, *Texas Forum of Teacher Education*, 3, 2-22.

RELIGIOUS PARTICIPATION AND ECONOMIC RECESSIONS

Reginald J. Harris, Georgia Regents University
Simon Medcalfe, Georgia Regents University

ABSTRACT

During the financial crisis which began in 2007, a number of articles explored the belief that religiosity increases during times of economic hardship. Using individual church data from over one thousand churches in the North Georgia Conference of the United Methodist Church from 1999 to 2009, this paper analyzes the relationship between religious participation and economic recessions. Participation is measured by attendance at the main Sunday worship service and the state of the economy is measured using various unemployment rates. Overall, recessions tend to be associated with declines in attendance at United Methodist Churches, although the effect is not strong.

Keywords: Recession, United Methodist Church, Religiosity, Church Attendance, Business Cycle, Economic Conditions, Church Growth.

INTRODUCTION

The last recession was particularly severe, being dubbed the great recession. It lasted 18 months from December 2007 until June 2009 with Gross Domestic Product declining about 5 percent. The national unemployment rate peaked at over 10 percent in October 2009. The S&P500 index fell over 56 percent between October 2007 and March 2009. The recession affected many individuals and firms adversely.

In the United States religious faith plays an important role in people's lives. According to the CIA World Factbook, 78.46% of Americans practice some form of Christianity. Protestants are 51.3% of the population (CIA 2007). Church participation affects ones sense of wellbeing and financial status as well as social, economic, and political behavior. Over 60 percent of the U.S. population are members of a church (Iannacconne, 1998). Over 40 percent of Americans report attending church weekly or almost weekly (Newport 2010). Hence, it is a significant research question to determine how church participation is being impacted by changes in the economy.

This paper analyses the effect of the past recession on church participation in Georgia. Using data from the North Georgia United Methodist Conference containing over one thousand churches, we analyze the effect of the business cycle on church attendance and find that the Georgia unemployment rate was negatively associated with church attendance.

LITERATURE REVIEW

Economic theory suggests there are two countervailing effects of the business cycle on church attendance. The substitution effect would suggest that church attendance increases in times of higher unemployment because the opportunity cost of time has decreased. Alternatively,

if church attendance is a normal good, declines in income associated with higher unemployment would be reflected in lower church attendance. Since theory is inconclusive on this matter, we review several empirical studies.

Azzi and Ehrenberg (1975) made the first systematic attempt by economists to analyze the determinants of religiosity. Religiosity is an individual's religious commitment and participation in church-related activities. They looked at how members of households allocate their time to religious involvement and how this varies with age. This is referred to as the "household life-cycle religious-participation profile". Two data sets were used, one using membership data and another using survey responses regarding church attendance. Azzi and Ehrenberg concluded that as real wages grow individuals are expected to shift to less time-intensive forms of religion. As unemployment rises, individuals are expected to shift to more time-intensive forms because the opportunity cost of participation declines. This could be an economic motivator for people to participate in religious activities during recessions.

Ellison (1991) noted that religion has been shown to give participants an increased sense of well being. He explored several possible motivating factors that may be responsible for providing that sense of well being. These factors included social integration, divine interaction, stress, and the effects of age and education. He found that attendance and private devotion contributes to well being, though indirectly, and that strong faith makes life trauma easier to withstand. Religion does give support for people in time of adversity. This would be consistent with there being an increase in participation during recessions when unemployment rises.

Beckworth (2009) explored how religiosity varies over the business cycle in Protestant churches. His research used data from three sources: data from a Pew Research Center survey administered during the 2001 recession period to assess weekly attendance, annual membership data spanning 1968-2004 to see if there is a relationship to NBER business cycle dates, and quarterly data from 1950:Q1 through 2006:Q4 to determine the relationships between Seventh-day Adventist converts and macroeconomic variables. He found that there is a difference in how the business cycle affects mainline churches versus how it affects evangelical churches. Growth for evangelical denominations was countercyclical while growth of mainline denominations had a procyclical component.

The differential effect on mainline and evangelical denominations may be explained by several factors. Individuals in mainline churches tend to be wealthier individuals whose income is more related to salaries as opposed to hourly wages. Whereas in an economic downturn, the hourly worker has his hours cut freeing time for religious participation, the salaried worker must maintain the same hours or perhaps work longer hours to compensate for diminished staff due to lay-offs. In an upswing in the economy, the hourly worker has the opportunity to benefit from overtime pay rates. It is to his advantage to work longer for more money. The salaried worker may find business easier to come by with less effort. This would allow the salaried worker more time to pursue religious activities. Moreover, higher income business professionals may benefit from the networking opportunities that church attendance may provide when the economy is doing well. Fellow attendees have business deals to make and opportunities they can share. When the economy is poor, there is less to be gained from networking since your fellow attendees do not have as many business deals to make or opportunities to share. Also, a church with accomplished fellow congregants may have a status conscious culture. During a downturn there could be a high social capital cost for being seen as suffering from the economy. This could reduce the incentive for attendance.

Other studies have also found a differential effect of the business cycle on church attendance that depends on the denomination. Sales (1972) focused on patterns of joining authoritarian and nonauthoritarian organizations using conversion rates. He found that individuals were drawn to authoritarian organizations during hard times, but that they gravitated to nonauthoritarian organizations when times were good. Authoritarian churches share similar characteristics such as the demand to absolute obedience to church leadership or to the Divine, condemnation of those who don't agree with their beliefs, literal interpretation of Scripture, strong concern about sin, and emphasis on mysticism. In contrast, nonauthoritarian churches tend to share other characteristics such as allowing parishioners more discretion in their decisions, assuming a more tolerant attitude to those who disagree, emphasizing a more intellectualized approach to scripture interpretation, and being less concerned over sin.

McCann (1999) examined the hypothesis that increased sense of threat is related to a stronger attraction to authoritarian churches and a weaker attraction to nonauthoritarian churches. McCann used membership data from two authoritarian and two nonauthoritarian denominations over the years from 1928-1986 to compare annual membership changes. He also studied twenty-five authoritarian and nonauthoritarian denominations during three distinct ten year periods from 1955-1985 that differed in threat levels. He found that during threatening economic times authoritarian denominations saw their greatest gains or smallest losses and nonauthoritarian churches saw their greatest losses or smallest gains during more threatening periods.

Overall, theory is inconclusive about the effects of a recession on church participation; the substitution effect suggests participation should increase because of the lower opportunity cost of time while the income effect suggests that participation should decrease. The few previous empirical studies have produced conflicting results; some suggest participation increase in times of economic stress and others reporting a decline. The effect depends on the denomination. This study is unique in investigating how religious participation and giving in a regional subset of a denomination are affected by local and national economic trends. No previous study has looked solely at the United Methodist denomination.

DATA, METHODOLOGY AND RESULTS

The North Georgia Conference of the United Methodist Church includes over one thousand individual churches. Data on church attendance is collected from statistical tables published by the conference. Attendance is recorded at the main Sunday service. Initially, the data is aggregated across all churches for the years 1999-2009. Our measures of the business cycle are the US unemployment rate and the Georgia unemployment rate; we also run our regressions with one year lags of each of these variables.

The results are presented in the first four rows of Table 1. The coefficients on the Georgia unemployment rate and the lagged unemployment rate are both negative but are not statistically significant. The coefficients on the US unemployment rate variables are similarly not significant. However, if churches serve a local community then a better measure of the state of the economy would be the county unemployment rate in which the church is located. The other advantage of using county unemployment rates is that is allows us to use individual church attendance data over the eleven years. Using a fixed effects model we find that an increase in the county unemployment rate decreases church attendance. Although this effect is statistically significant at the 5 percent level, the magnitude of the effect is small: A one percentage point increase in the county unemployment rate decreases church attendance by less than one person! The effect is of

similar magnitude using either the county unemployment rate of the lagged county unemployment rate. The average attendance at the churches in the dataset is 142 so a one percentage point increase in the unemployment rate is associated with a 0.6 percentage point decline in attendance. In other words, we could say that UMC church attendance is inelastic.

For all the regressions the R-squared value is very low (under 10 percent and sometimes almost zero). This is not surprising with the simple regression equation we have used. However, it does suggest that factors other than the economy are more important determinants of church attendance. For example, the spiritual health of the church, the role of the pastoral staff, and other non-economic variables seem likely to determine attendance. Future research should endeavor to include some of these other variables in their analysis.

CONCLUSION

The last recession was particularly severe affecting many individuals and organizations. One understudied organization is the church and church participation even though religious life is important to many people in the US. We found that attendance at UMC worship services declined during periods of high unemployment over the last decade or so. This is in line with previous studies that found nonauthoritarian churches tended to see attendance decline in recessions (Beckworth, 2009; Sales, 1972; and McCann, 1999). However, the effect is rather small. If local unemployment increased by five percentage points this would only result in a decrease of four people in attendance at the main worship service. Overall, despite the severity of the last recession the effect on religious participation at the United Methodist Church in North Georgia was slight.

REFERENCES

Azzi, C and Ehrenberg, R. (1975). "Household Allocation of Time and Church Attendance." *Journal of Political Economy*. 83(1): 27-56

Beckworth, D. (2009). "Praying for a Recession: The Business Cycle and Protestant Religiosity in the United States." Available at SSRN: http://ssrn.com/abstract=1103142 or http://dx.doi.org/10.2139/ssrn.1103142

Ellison, C. (1991). "Religious Involvement and Subjective Well-Being." *Journal of Health and Social Behavior*. 32(March): 80-99.

Iannaccone, L. (1998). "Introduction to the Economics of Religion." *Journal of Economic Literature*. 36(3): 1465-1494

McCann, J.H. (1999). "Threatening Times and Fluctuations in American Church Memberships." *Personality and Social Psychology Bulletin*. 25(3): 325-336.

Sales, S. (1972). "Economic Threat as a Determinant of Conversion Rates in Authoritarian and Nonauthoritarian Churches." *Journal of Personality and Social Psychology*. 23(3): 420-428

Vitello, P. (2008). "Bad Times Draw Bigger Crowds to Church." *New York Times*. December 14:A1.

Table 1: Effect of business cycle on participation and giving

	Dependent variable: Attendance
GA unemployment rate	- 448 (0.37) R^2=.09 N = 11
GA unemployment rate (lagged one year)	- 524 (0.65) R^2 = .03 N = 10
US Unemployment rate	- 341 (0.56) R^2=.04 N = 11
US Unemployment rate (lagged one year)	690 (.57) R^2 = .04 N = 10
County unemployment rate	-.793 (0.03) R^2 = 0.004 N = 9919
County unemployment rate (lagged one year)	-.762 (0.03) R^2 = 0.004 N = 9916

Note: p-values in parentheses

COMPENSATION DISCRIMINATION IN THE NATIONAL FOOTBALL LEAGUE

Nancy J. Burnett, University of Wisconsin Oshkosh
Lee J. Van Scyoc, University of Wisconsin Oshkosh

ABSTRACT

Keefer's recent article in the Journal of Sports Economics (2013), "Compensation discrimination for defensive players: applying quantile regression to the National Football League market for linebackers," finds wage discrimination in the National Football League (NFL) for linebackers. We examine the market for NFL tight ends using the same techniques as Keefer, though we explore only rookies rather than all current players and tight ends rather than linebackers. While we would expect to find stronger evidence of discrimination in the rookie market, as rookies are captured sellers, we find no pervasive pattern of pay discrimination by race for tight ends.

INTRODUCTION

The study of racial discrimination in professional sports has been an actively studied topic for decades, though much of this work has been done in sports other than the National Football League (NFL). However, the NFL itself has become highly conscious of discrimination in both hiring and compensation by race since the early 2000's. The NFL created a diversity committee in October of 2002 expressly to hunt for discriminatory practices and offer recommendations to reduce racial discrimination of all types in the league.

Recommendations from the diversity committee included training and development programs aimed expressly at minorities. For instance, one recommendation, the so-called 'Rooney Rule' named after the chair of the committee Dan Rooney, ensures that at least one minority coach must be considered for every high level coaching position that opens in the league. The opportunity to discriminate against either coaches or players, however, cannot be fully eradicated by such rules. Though this paper does not deal with coaches, there exists the possibility of compensation discrimination by race as long as there is the possibility of collusion again black players or coaches. Without collusion, teams not discriminating would be able to hire talented black players at lower cost and increase their winning record, effectively bidding out most of the discrimination among teams (assuming teams put winning as a top priority). The incentive to break collusion of this type is strong, however, because even a single extraordinarily talented player can make a huge impact on a team.

When it comes to discrimination, Becker (1973) discusses the three classic types or 'tastes' of discrimination. They are employer discrimination, employee discrimination, and customer discrimination. Employer discrimination represents the situation where an employer simply has a preference for hiring one type of employee and pays that group more than the non-favored group. The second type of discrimination mentioned by Becker, that of employee discrimination, is more subtle. In that case, one group of employees displays such a strong preference not to work with someone from a non-favored group that the resultant discord essentially forces the employer to discriminate to prevent the loss of overall productivity from

attempting to mix the groups. The third form of discrimination is customer discrimination, in which customers have such strong preferences that they that they change their purchasing habits significantly enough to provide an incentive for employers to discriminate against the non-favored group in order to maintain market share and profits.

Discrimination should be non-profit maximizing as a rival could hire the non-favored employees and create higher profits (in this case team wins, likely driving team revenue and profit), presumably driving the discriminating producer out of business. There have been suggestions, however, that discriminatory practices may not result in a negative outcome for the discriminators fast enough to bid it out of the market rapidly (Hellerstein, Neumark & Troske, 2002).

Professional sports, like any other for-profit business, define success by making profits. One way to make profit is to have a winning season. Having the best players is one way to attempt to have a winning record. Indeed, an argument can be made that in sports even one player can make a difference in a season record, a much stronger impact than a single employee is likely to make for a firm in other output markets. Therefore, there is a strong incentive for teams to find and retain the best talent available, regardless of race. This reduces the likelihood of employer discrimination in a for-profit scenario where profit hinges on having talented players. A team's talent scout or general manager who does not accomplish this will not retain his job. Should a team discriminate by pay, the rookie experiencing such discrimination will be highly unlikely to remain with that team after his initial contract expires and he becomes a free agent, so that discriminatory teams will be unable to retain the best talent. That incentive will lead to stronger incentives in sports against discriminatory actions than in the general market. It also suggests that the market for rookie players is most likely to have this type of discrimination, as a free agent player will have market forces to drive wages toward equilibrium.

Employee discrimination, when employees reduce overall productivity because they so dislike working with a fellow employee, can cause an employer to discriminate so as to prevent the reduction in overall productivity. While it is possible that players could find a fellow player so disagreeable that team cohesion would be reduced sufficiently so that a team may wish to reduce pay or eliminate that player, it is highly unlikely that this would be identifiable for a rookie player at least based on an individual's personality. A team with significant racist tendencies may attempt to ostracize all players of the other race to the extent that a team would recruit with such tendencies in mind; however, if a team were to be so racist, one would expect to see such a team already exhibit race bias in previous player choices. Such teams would already be nearly all white or all black, in other words. The fact that we do not currently see teams with such bias in membership suggests that this type of racism is not terribly prevalent.

Customer discrimination, another of the classical 'tastes' for discrimination, may play a more significant role in sports. Customers can engender discrimination under certain circumstances, particularly where there are 'fan favourite' players who have enough fan loyalty to affect profits with stadium attendance, team memorabilia sales, or even TV market share. When players of this type come up for a new contract a team may have the incentive to overpay these players relative to their contribution to the team's win/loss record. For instance, consider a player of average skill (or one whose skills have diminished due to age or injury). Logically that player would likely either be released or see a compensation reduction. However, if that player is a fan favourite, the team may want to appease those fans by resigning such a player, even potentially paying higher compensation than he is objectively worth. It has been shown multiple times that fans do seem to suffer from race preference (see Burnett & Van Scyoc, 2004, among

many others). If this is the case, popular players may be overpaid on subsequent (non-rookie) contracts. A previous study (Keefer, 2013) uses exclusively non-rookie players who may be subject to this type of discrimination. Our study uses only rookie players, making this form of discrimination unlikely to affect our results, as fans have yet to establish strong favourites among pre-professional players.

Academic work has been done on salary discrimination by race among players of the NFL. Mogul (1973, 1981) did not show difference between black and white players' salaries. The samples involved, however, were small and were collected by survey, bringing in the potential for response bias. Another fairly dated study by Kahn (1992), using data from the 1989 NFL Players Association, found that white players earned about 4.1% more than non-whites, though this difference was not statistically significant, supporting Mogul's (1973, 1981) conclusions. Gius and Johnson (2000) used a data set of 938 NFL players from the 1995 season. Contrary to previous studies, they found that white players made 10% *less* than black players, even when controlling for player position. Berri and Simms (2009) looked exclusively at the quarterback position, though that position is staffed mostly by white players. They found bias in favour of white players, perhaps because most black quarterbacks rely more on rushing than white quarterbacks, which is not usually rewarded as well as passing. Keefer (2013) studied 1,575 linebackers, between 2001-09, using various measures of quality and found salary discrimination against black linebackers. As his work represents all linebackers including those who have been in the league long enough to have become free (or restricted) agents and gone through multiple contract negotiation processes, he may have picked up issues involving customer discrimination as well as forms of employee and/or employer discrimination.

OUR APPROACH

Our sample of rookie players in the wide receiver position over the 2000-09 period (the only period for which data is available and consistent) represents a 'captive' market since the player is "forced" to sign with the team that drafted him. If he does not sign with his drafting team, he must sit out for a year which represents a very large opportunity cost. The result is a market that is less than fully efficient. This problem does not dissipate until competition is brought to bear on a player's salary. In this case, after a player has been in the league for three seasons and completed his initial contract, a player becomes a restricted free agent and is able to receive bids from other clubs (the current team has the right to match any offers to a restricted free agent in order to retain the player, however). After four or more seasons and completing his initial contract, a player can become an unrestricted free agent so that he can sign with any club, so the market becomes fully competitive for those players. In either case, the market setting a player's salary after the initial contract will be more efficient than the draft market as competition is allowed to operate to at least some extent, likely bidding away discrimination. Therefore, we have restricted our study to rookies with the idea that we will be more likely to detect generalized racial discrimination in this market. Individual players may be paid at a rate not commensurate with their measurable talents even in the case of free agents for non-race related reasons such as being a fan favorite, though this is not likely to be the case with rookies. One drawback to this approach is that we are unable to use quality measurements obtained from play within the NFL as rookies have yet to play at that level.

We concentrate on the tight end position. This position has players of both races, with about 44.5% of the players being black. Further, there are always a fairly large number of players

in this position in the rookie draft and as teams generally have an ongoing need for tight ends, the draft is fairly active for this position.

MODEL AND DATA

Model

We begin with the standard ordinary least squares (OLS) earnings function for player *i*'s salary for year *t*, Y_{it}, with a vector of independent variables, including a racial identifier, in $x_{i,t}$:

$$Y_{i,t} = \alpha + \beta x_{i,t} + \varepsilon_{i,t}$$

Equation 1: OLS Earning Function, which includes a racial identifier.

The traditional OLS earnings function approach estimates parameters at the conditional mean and is highly efficient but is quite sensitive to outlier values. The case at hand, athlete salaries, is one that is particularly prone to outliers. We follow Keefer (2013) and others, in expanding the analysis to the use of quantile regressions (see Koenker and Bassett, 1978), which is far more robust to the presence of outliers and non-normal distributions. Quantile regressions, whether segregated into quartiles, quintiles or other grouping, allows for the effect of the independent variables to vary across the distribution, with the assumption that the conditional θ^{th} segment of the dependent variable is a linear combination of the independent variables. There are several examples of quantile regression in the sports labor market, including Keefer (2013), Burnett and Van Scyoc (*forthcoming*) and Vincent and Eastman (2009).

Quintile regressions subdivide the data into five sub-groups based on the dependent variable. In our case, we follow Keefer (2013), and others, and use sub-groups of the lowest 10%, up to 25%, median, upper 25%, and top 10% (listed as 10%, 25%, 50%, 75%, and 90%). Once the data has been broken up in that way, OLS regressions are run (including a racial identifier) for each sub group. That way, the coefficients are determined not relative to the overall mean of the dataset but for the mean of each sub group. Discrimination is found if there are significant differences by race in each group. As the top and bottom groups will likely have the biggest outliers, we are able to reduce the outlier effect on the remaining results.

A second method of exploring data for discrimination, involves breaking the data into two groups by a common characteristic (such as race), then estimating each group separately. A statistical comparison of those two groups is performed for overall differences, then differences due to endowment or characteristic differences (such as player quality variables), and then by coefficient on the contribution of those characteristics to the total differential. If there are any significant differences in the coefficient estimates between the racial groups, that indicates quality is paid differently dependent upon race and discrimination is identified. This is known as the Oaxaca-Blinder decomposition. In general, using H as the designator for the higher paid group and L for the lower paid group, and the various vectors of income (Y_{it}) and descriptive variables ($x_{i,t}$, less the racial descriptive), our model becomes:

$$Y_{i,t}^{H} = \alpha^{H} + \beta^{H} \times x_{i,t}^{H} + \varepsilon_{i,t}^{H}$$
$$Y_{i,t}^{L} = \alpha^{L} + \beta^{L} \times x_{i,t}^{L} + \varepsilon_{i,t}^{L}$$
$$\bar{Y}^{H} = \alpha^{H} + \beta^{H} \times \bar{x}^{H}$$

$$\bar{Y}^L = \alpha^L + \beta^L \times \bar{x}^L$$
$$\bar{Y}^H - \bar{Y}^L = (\alpha^H - \alpha^L) + (\beta^H - \beta^L) \times \bar{x}^L + \beta^H \times (\bar{x}^H - \bar{x}^L)$$

Equation 2: Oaxcaca Blinder Regression Decomposition, where H and L designate the higher and lower paid group and where the coefficient differences, $(\alpha^H - \alpha^L) + (\beta^H - \beta^L) \times \bar{x}^L$ are the measure of discrimination.

The Oaxaca Blinder decomposition uses OLS estimates at the conditional mean, however, we extend this decomposition to the quantile regression results rather than being limited to the overall conditional mean, due to the impact of outlier data inherent in athlete salaries. Essentially, we substitute the segmented function for salary for each group, separated by a race binary variable **Black** (1 for black and 0 for white). This specification allows the estimator to be used for hypothesis testing and inference (see Melly 2006, and Keefer 2013). Bootstrapping is necessary to determine the estimated standard errors.

$$F^{-1}_{Y(0)}(\theta) - F^{-1}_{Y(1)}(\theta)$$
$$F^{-1}_{Y(0)}(\theta|\mathbf{Black} = 1) - F^{-1}_{Y(1)}(\theta|\mathbf{Black} = 1)$$

Equation 3: Quantile Treatment Effects (QTE), where $F^{-1}_{Y(\mathbf{Black})}(\theta)$ is the Θ^{th} segmented function of *T* for **Blacks**.

Data

Our sample consists of an original data set of 281 players that joined the NFL in the position of tight end during the 2000-2009 seasons (there are a small number of omitted players for whom we could find no pictures or mention of racial identifier). USA Today maintained an online database of player salaries for all major sports and maintained consistency in their recording methods only between the years of 2000 and 2009 (they changed collection methods before that time and report no individual data after 2009). The database provides several different measures of player income, including Base Salary, Cap Value (the portion of a players contracted salary that contributes to the team's salary cap) and Total Salary (Base Salary including all Bonuses). We take the natural logarithm of Total Salary, adjust for inflation, and obtain **Ln Real Salary** for each player in our dataset (substituting other measures of salary, such as Cap Value, did not significantly affect our results). Even though the rates of inflation during this period (2000-09) were quite low and fairly stable, we use these inflation adjusted salary values as our dependent variable (the authors also tried other variations of salary data – using Base Salary alone, with and without other bonuses, with and without inflation adjustment and find no significant differences in results).

We measure player quality using draft pick order and status. We believe that major league general managers and scouts, for whom judging quality is an integral part of their job, are likely to have a better idea of quality than any mere compilation of barely comparable (or completely non-comparable) statistics. Certainly, teams try to take into account intangibles as much as possible, even though such characteristics do not show in any measurable variable. College playing statistics are non-comparable due to variances across the wide array of college teams. Not all players even attend the NFL combine before the draft, and even those that do attend do not all compete with the same tests, making those scores also non-comparable.

Our first variable, **Draft,** is the overall draft number regardless of round, for example, the 8th player taken would have a draft order of 8, even though he may be the first tight end player taken. Draft number will vary by quality as well as by need, though teams have an ongoing need for tight ends (unlike some other positions) so they are likely to draft a high quality player even if that position is *not* a team's first priority. Player quality is measured solely by draft pick number for those who were drafted (**Draft**) with low numbers signifying better draft picks and, presumably, higher quality. Those players who were not drafted were assigned a draft pick number of 0, which simply separates out those players who were not drafted and generates an intercept shift for those players. A binary variable then designates those players that ultimately received a contract without being drafted picks up those players (**Undraft**). This process generates estimated salaries that decline with increased draft pick number and greatly decreased estimated salaries for undrafted players. An example of this process shows predicted salary levels, using the simple OLS model (see below). We generated variables for three hypothetical players, all with the same racial identifier (**Black**=0, meaning a white player), but with different draft numbers and status (**Draft**=1, **Draft**=100, and **Undraft**=1) assuming the year is 2005 (approximately half way through the timespan of our data). We see the following results:

> For a white first pick (best, meaning lowest, draft number) we would expect an **Ln Real Salary** = 9.084221 + 0.1373058 (**Black**=0) - 0. 0101995 (**Draft**=1) - 1.895866 (**Undraft**=0) = 9.0740215 , for a nominal estimated total salary of $1,658,723.
> For a white player with a draft pick of 100 we would expect **Ln Real Salary =** 9.084221 + 0.1373058 (**Black**=0) - 0. 0101995 (**Draft**=100) - 1.895866 (**Undraft**=0) = 8.064271 , for a nominal estimated total salary of $620,827.06.
> For a white player that was not drafted, but merely signed at some point after the draft we would expect an **Ln Real Salary =** 9.084221 + 0.1373058 (**Black**=0) - 0. 0101995 (**Draft**=0) - 1.895866 (**Undraft**=1) = 7.188355 , for a nominal estimated total salary of $258,562.27.

These results indicate, as we would expect, that the better (lower) the draft pick, the higher the expected salary and undrafted players would see the lowest expected salary.

General summary statistics by race are found in Table 1. We see that black players appear to be paid, on average, quite a bit more than white players. Further, we note that the average draft number (of those players who were drafted) was higher (worse) for whites than it was for blacks, which may explain the pay differential. However, draft pick number and status are based upon team preferences and if those preferences are discriminatory, draft numbers may simply reflect the racist views of the league. If that is the case a single *non*-discriminating team would appear to be able to pick up a player whose talent is under estimated (or under-appreciated due to the race of the player) for less salary than he would otherwise be worth. Should that be the case, the non-discriminating team would get more talent for less overall pay than other teams. That should prove significant incentive to break any collusion for race based pay discrimination. Additionally, we see slightly more black players were undrafted than white players, suggesting that teams were not completely passing over white players in favor of black players.

Table 1 Summary Statistics (Means) by Race			
	All	Black	White
Total Salary	$620,873.90	$674,964.40	$577,532.20
Black	0.4448		
Undrafted	0.5338078	0.5440	0.525641
Draft (of those drafted)	130.458	125.0351	134.6351
N	281	125	156

A further examination into salaries by race and draft status is found in Table 2. We continue to see that black players are paid more than white players and that pay differential becomes much larger among drafted players than among undrafted players. Among the undrafted players pay is far more equal, though there is, again, a slight positive differential in favor of black players.

Table 2 Mean Total Salary by Draft Status and Race		
	Black	White
All	$674,964.4 (125)	$577,532.2 (156)
Drafted	$934,909.7 (57)	$749,559.9 (74)
UnDrafted	$457,069.0 (68)	$422,287.6 (82)

Counts in parentheses.

ESTIMATION RESULTS

Examining the OLS initial results using a dummy variable as a racial indicator, we find the binary racial indicator variable, **Black**, is significant at the 20% level and positive, suggesting that blacks are paid better than whites overall. The quantile approach, as pioneered by Melly (2006) does not show that blacks are paid better until reaching the top 10% of players when results are estimated by sub-groups, and that difference is only significant at the 20% level as well, suggesting that outlier salaries may be driving the overall OLS results. In all of these

income groupings (lowest 10%, lowest 25%, median group, top 25% and top 10%) it is clear that the other characteristics of player quality (draft number and draft status) holds more explanatory power for salary than race.

One characteristic of the traditional OLS method is that it is quite sensitive to outliers, a factor inherent in athlete salaries. Therefore, the very nature of this type of dataset would suggest strongly that traditional OLS methodology would be less than optimal. Using the quantile approach, so that outliers do not shift the overall results, creates a far more robust estimation method. In this type of estimation, outliers only skew the bottom and top sub-groupings. Indeed, in our results we see that race only shows significance (and then only at the 20% significance level) in the top group which is potentially prone to the outlier effect.

Table 3
OLS and Quantile Estimates for Dummy Variable Regressions

Dependent Variable =Ln Real Total Salary						
		Quantile				
Variable	OLS	Q10	Q25	Q50	Q75	Q90
Black	0.1373058* (0.0916631)	0.2638502 (-0.009608)	-0.0148139 (0.0657167)	0.0503832 (0.0743598)	0.1216598 (0.1260055)	0.2296609* (0.1420985)
Draft	-0.0101995*** (0.0009509)	-0.009608*** (0.0029646)	-0.0096281*** (0.0010964)	-0.0090869*** (0.0008266)	-0.0093056*** (0.0009613)	-0.0115692*** (0.000923)
Undraft	-1.895866*** (0.153916)	-1.968088*** (0.4769316)	-2.020471*** (0.1486877)	-1.85943*** (0.1426745)	-1.713681*** (0.1828468)	-1.662762*** (0.2718102)
Constant	9.084221*** (0.1479093)	8.9091122*** (0.3044756)	8.874375*** (0.1477246)	9.015474*** (0.1265668)	9.294398*** (0.1497514)	9.214548*** (0.25395)
R^2 (pseudo)	0.3619	0.0898 (pseudo)	0.1873 (pseudo)	0.2789 (pseudo)	0.3014 (pseudo)	0.2947 (pseudo)
Observations	281					

Note. Standard errors in parentheses. Quantile standard errors computed from 1,000 bootstraps. R^2 reported for OLS, Psuedo R^2 reported for quantile regressions.
*Significant at 20%.
**Significant at 5%.
***Significant at 1%.

Table 4 shows results from the Blinder-Oaxaca decomposition method, which breaks down the data set into high and low paid groups and then compares regression results from the two groups (either from the overall dataset or the quantile sub-groups). It is in the comparison of the two groups that discrimination can be seen if differences arise from different estimates of the coefficients rather than from different endowments (or characteristics). In the two groups, one would expect that the higher paid group would have better measured characteristics (in this case, we would expect higher paid athletes to have been drafted, rather than undrafted, and have lower

draft pick numbers. As long as differences across the two groups can be accounted for with such measurable characteristic differences, there would be no evidence of discrimination. It is if the two groups *also* have different estimated coefficients that we would be seeing discrimination. For example, that would be the case if one group were paid an average of $50,000 for each incremental better draft pick, while for the other group the players were paid an average of $20,000 for each incremental better draft pick. This method, then, splits the data set between the two racial groups, setting whites as Group 1 and blacks as Groups 2 and compares the resultant differences, so that positive differences indicate higher values for whites and negative differences show higher values for blacks. Discrimination, or differences according to group status, is seen when there is significance on the coefficient differences between the groups suggesting that players with identical characteristics are paid differently based only on their race. Differences to endowments or overall differential can be attributed to differences that are may be due to the quantifiable characteristics of the groups (for instance, since blacks have better draft numbers and we would *expect* to see players with lower draft numbers get paid better, to the extent that blacks with lower draft numbers are paid better is not due to discrimination). What we see is that, for these overall groupings, there are negative values showing higher value for blacks for total differential, endowments, and even for the coefficient estimates. The differences are only significant at the 20% level for overall differences and coefficients. Since these results are not broken down by quintile they may reflect outliers that overshadow the results. Indeed, when looking at the QTE results of this decomposition method, it becomes clear that in no case are there significant differences between coefficients across the higher and lower groups within the quintile sub-groups.

Table 4
Decomposition Results, Oaxaca-Blinder and QTE

Dependent Variable =Ln Real Total Salary						
	Oaxaca-Blinder	QTE				
	OLS	Q10	Q25	Q50	Q75	Q90
Total Differential	-0.1723601* (0.1125624)	0.203304** (0.101427)	0 (0.054332)	-0.010831 (0.127886)	-0.240324* (0.129338)	0.050901 (0.127282)
Endowments	-0.0353832 (0.0667955)	0 (0.188346)	-0.013554 (0.043622)	0.030629 (0.0943)	0.11159 (0.118754)	0.072988 (0.119991)
Coefficients	-0.1373646* (0.089948)	0.203304 (0.180736)	0.013554 (0.037085)	-0.04146 (0.084984)	0.128734 (0.118528)	-0.022087 (0.125282)
R^2	Group 1: 0.3138 (White) Group 2: 0.4282 (Black)					
N	Group 1: 156 (White) Group 2: 125 (Black)					

Note. Standard errors in parentheses. Quantile standard errors computed from 50 bootstraps. R^2 reported for OLS Oaxaca-Blinder only.
*Significant at 20%.
**Significant at 5%.
***Significant at 1%.

CONCLUDING REMARKS

We find no body of evidence in our sample that there exists broad based racial salary discrimination in the position of tight end in the NFL between the years of 2000 and 2009. Traditional OLS results suggest that there may be such discrimination, albeit in favour of blacks; however, upon disaggregation of the data into sub-groupings that evidence evaporates. Likely the naïve estimates from the OLS model are unduly influenced by outlier data. When we look at the Oaxaca-Blinder style of estimation the evidence suggests that most of the difference in salary arises from differences in characteristics of the players with black players having better draft position and more of the white players being of the undrafted (and hence lower paid) variety. There is only one two sub-grouping that shows any slight evidence of discrimination and that is the top sub group of the OLS/Melly(2006) QTE. That evidence is only at the 20% significance level, and could well be due to a few outliers (specifically, outliers that would be a the top of the drat pick or 'stars' that could well be overpaid due to the hype of the draft). In the Oaxaca-Blinder decomposition method again we see significance at the 20% level for the overall group for different coefficients but that disappears in all of the Oaxaa-Blinder quintile analysis. In

general, what slight evidence we see in favour of discrimination shows first that blacks are *better* paid than whites and second is likely to be true only because of a few outliers from the very top of the drafted players.

This result runs contrary to Keefer (2013) but supports that from Burnett and Van Scyoc (*forthcoming*) and earlier results by Mogul (1973, 1981) and Kahn (1992). This, despite the fact that we used rookie data where we would expect to find the strongest evidence of such discrimination since we are dealing with captured sellers. There are several potential explanations why Keefer (2013) found discrimination while we did not. For instance, much of our data reflects players (and their salaries) who joined the league after the initial push in the early 2000's in the NFL to uncover and remedy discrimination, while the majority of the players in Keefer's data had been hired into the league before that time as he was working with all currently active players over those years.

Additionally, it is possible that black players actually have different skill sets, making them appear more talented based on the type of measured characteristics used by Keefer (2013), suggesting they would warrant higher salaries making it appear that there were actually discriminatory practices occurring. However, if pay is actually a reflection of overall ability (including intangibles that would not be picked up by the type of characteristics used by Keefer, but taken into account by NFL scouts and general managers and therefore reflected in draft pick numbers) and if it is the case that black players have more of these intangibles, pay may actually be appropriately allocated (cases of this would be shown with differences in endowments in the Oaxaca-Blinder and QTE decompositions like we observed).

Also, for those players in the draft, rather than for undrafted players, there is also some limit on the variability of salary offers relative to draft pick rankings. Salary and contract offers for those players are closely scrutinized and it is rare for salaries to 'overlap' draft pick ratings (for instance, the player picked second or third would not usually be paid more than the player picked first). Hence, for drafted players there is a sort of built in pay scale that does not allow much leeway for discrimination. If drafted players are simply paid much more than undrafted players, there could be some discrimination if, say, black players are more likely to be drafted and white players are more likely to be signed as undrafted players. In our dataset, however, more of the undrafted players are black rather than white (though this is by a slim margin). Therefore, we again conclude that we see no overwhelming evidence of racial discrimination in pay in the position of tight end in the NFL incoming players over the years 2000-2009.

REFERENCES

Becker, G. (1973) The Theory of Discrimination, in *Discrimination in Labor Markets,* Ashenfelter and Rees, eds., (Princeton NJ: Princeton University Press), 3-33.

Berri, D. J. and Sims, R. (2009) Race and evaluation of signal callers in the National Football League, *Journal of Sports Economics*, **10(23)**, 23-43.

Burnett, N. J. and L. J. Van Scyoc (*forthcoming*) Compensation discrimination for defensive players: applying quantile regression to the National Football League market for linebackers and offensive linemen, *Journal of Sports Economics.*

Burnett, N. J. and L. J. Van Scyoc (2004) Discriminating Buyers of Baseball Cards: Does Race Affect Value?, *Journal of Economics and Economic Education Research*, 5(3), 101-110.

Gius, M. and Johnson, D. (2000) Race and compensation in professional football, *Applied Economics Letters*, **7(2)**, 73-75.

Hellerstein, J., D. Neumark, and K. Troske (2002) Market Forces and Sex Discrimination, *The Journal of Human Resources*, **37(2)**, 353-380.

Kahn, L. (1992) The effects of race on professional football players' compensation, *Industrial and Labor Relations Review*, **45(2)**, 295-310.

Keefer, Q. A. W. (2013) Compensation discrimination for defensive players: applying quantile regression to the National Football League market for linebackers, *Journal of Sports Economics*, February 14(1), 23-44.

Koenker, R. & Bassett, G. (1978). Regression Quantiles. *Econometrica, 46,* 33-50.

Melly, B. (2006). *Applied quantile regression*. (Doctoral dissertation). University of St. Gallen, Switzerland.

Mogul, R. G. (1973) Football salaries and race: some empirical evidence, *Industrial Relations*, **12(1)**, 109-112.

Mogul, R. G. (1981) Salary discrimination in professional sports, *Atlantic Economic Journal*, **9(3)**, 106-110.

Vincent, C., & Eastman, B. (2009). Determinants of pay on the NHL: A quantile regression approach. *Journal of Sports Economics, 10,* 256-277.

THE IMPACT OF DIFFERENT AGES AND RACE ON THE SOCIAL SECURITY EARLY RETIREMENT DECISION FOR MARRIED COUPLES

Diane Scott Docking, Northern Illinois University
Rich Fortin, New Mexico State University
Stuart Michelson, Stetson University

ABSTRACT

The purpose of this study is to examine the impact of age differences on the social security early and delayed retirement decision for married couples. This paper extends the analysis of Docking et. al. (2013) to couples of different ages. This analysis is done for married couples by race. More specifically, we analyze the 9 married couple combinations for the following races: Whites (W), Hispanics (H) and Blacks (B). The nine husband/wife combinations are: WW, BB, HH, WB, BW, WH, HW, BH and HB. We develop an Excel model to compute the breakeven IRR for each of the 9 race combinations. Following Blanchett (2013), three claiming scenarios are considered: receiving benefits early (e.g., at age 62 versus 66); the maximum realistic delay period (e.g., at age 62 versus 70) and delaying benefits past full retirement age (e.g., age 66 versus 70). Within these 3 claiming scenarios we examine couples by race combination who retire at the same age with age differences of 0, 4, 7 and 10 years. We assume the non-working wife (female) is younger than the working husband (male). The breakeven IRR's can be interpreted as follows: If a couple's opportunity cost of capital (which can be considered a hurdle rate) is greater than (less than) the computed breakeven IRR, the couple should retire at the earlier (later) age. For the age 62 versus 66 comparisons the BE IRR's uniformly decrease as the age difference increases. Since, as noted above, these IRR's are hurdle rates, this implies that greater age difference couples should retire earlier since the hurdle rate is less to overcome than at a smaller age difference. These results should be interpreted with caution however since an inflection point occurs at the age 62 versus 67 comparison (not shown in our tables) and continues onto the age 62 versus 70 comparison where the IRR's uniformly increase with age differences. We attribute this inflection point to the interaction of an increasing time gap between the early and delayed retirement with a constant set of age differentials. This age 62 versus 70 comparison implies that greater age differences involve a greater hurdle and the smaller the age difference the greater the incentive to retire earlier since the hurdle rate is lower. The results for the age 66 versus 70 comparison are similar to the age 62 to 70 comparison with the breakeven IRR's increasing with age differences although the numbers themselves are quite small by comparison and would seem to suggest early retirement at all age differences given the low hurdle rates to overcome. We also examine breakeven IRR's for couples by race combination who retire at different ages and who have a positive age difference. More specifically, we examine the impact of age differences on an early male/female retirement of 66 and 62 respectively versus a late male/female retirement of 70 and 66 respectively. In all 9 race combinations the breakeven IRR's decline as the age differences

increase. This suggests that the greater the age difference the greater the incentive to retire early as the hurdle rate is lower to overcome.

INTRODUCTION

The purpose of this study is to examine the impact of age differences on the social security early and delayed retirement decision for married couples. This paper extends the analysis of Docking, Fortin and Michelson (2013) to couples of different ages. This analysis is done for married couples by race. More specifically, we analyze the 9 married couple combinations for the following races: Whites (W), Hispanics (H) and Blacks (B). The nine husband/wife combinations are: WW, BB, HH, WB, BW, WH, HW, BH and HB. We develop an Excel model to compute the breakeven IRR for each of the 9 race combinations. Following Blanchett (2013), three claiming scenarios are considered: receiving benefits early (e.g., at age 62 versus 66), the maximum realistic delay period (e.g., at age 62 versus 70), and delaying benefits past full retirement age (e.g., age 66 versus 70). Within these 3 claiming scenarios we examine couples by race combination who retire at the same age with age differences of 0, 4, 7 and 10 years. We assume the working husband (male) is older than the non-working wife (female). We also examine a scenario where the couples retire at different ages with positive age differentials of 4, 7 and 10 years.

Individuals born between 1946 and 1954 can retire with full social security benefits at their full retirement age (FRA) of 66. The FRA gradually rises until it reaches 67 for people born in 1960 or later. However, individuals have the option to retire earlier or later than their FRA. The earliest one can retire is age 62, and the latest is age 70. Early retirement is attractive for many reasons: social security benefits (SSB) and rules can change, health concerns, and increased demand for leisure. However, SSB are permanently reduced by an actuarial reduction factor (5/9 of 1% for the first 36 months and 5/12 of 1% per month thereafter for early retirement). Delayed retirement is attractive because SSB are increased by a delayed retirement credit (DRC) of 8% for each year of delay after FRA up to age 70.

The results of previous research into the social security early and delayed retirement decision for married couples have been mixed. This paper extends the analysis of these prior studies to examine the role that race and age differences between spouses have on their retirement decisions. We will create a spreadsheet to model various retirement scenarios that will be beneficial for individual investors and their advisors.

LITERATURE REVIEW

There has been an extensive number of studies on the early versus delayed social security retirement decision for married couples although none have explicitly addressed the age difference issue across race categories as this study does. For a review of prior literature, see Docking, Fortin and Michelson (2012, 2013).

Only a few studies have looked at the age difference between the spouses in determining the optimal retirement age. Coile, Diamond, Gruber and Jousten (2002) find that if the husband is older than the wife, then he should delay retirement to age 65; but if the wife is 5 years older than

her husband, he should retire early at age 62. Munnell and Soto (2007) show that as the age difference between the spouses (husband minus wife) increase, the wife should claim earlier (age 62) and the husband should claim later (age 69). Sun and Webb (2009) show that if the wife is 3 or more years older than her husband, he should retire at 69 and she at 66. Tucker (2009) says both should retire at age 62 no matter the age difference. McCormack and Perdue (2006) assume the husband is 7 years older than his wife and the husband has the higher earnings. They conclude that both should retire at age 62.

Docking, Fortin and Michelson (2013) look at the impact of race on the retirement decision for married couples of the same age. They compute a breakeven (BE) internal rate of return (IRR) for each of nine race combinations from age 62 through age 70. The greater the BE IRR, the more optimal for a couple to retire later. Results are fairly uniformly consistent across the nine race combinations: BE IRRs for a given base age are, in general, monotonically decreasing compared with older ages. The highest BE IRRs are for couples with a Hispanic husband, and the lowest BE IRRs are for couples with a White husband. This paper will expand on the Docking et. al. (2013) study and explore the effect of race and age difference on the retirement decision of married couples.

HOW SOCIAL SECURITY WORKS

A detailed description of how social security works can be found in Docking, Fortin and Michelson (2012, 2013). Briefly, individuals aged 62 or older who had earned income that was subject to the Social Security payroll tax for at least 10 years (40 quarters) since 1951 are eligible for retirement benefits.

No matter what your FRA is, you may start receiving benefits as early as age 62. However, if you start your benefits early, they will be reduced a fraction of a percent for each month before your FRA. This reduction is permanent. A worker with a FRA of 66 who claims early at age 62 receives 75% of their FRA benefit amount; a worker with a FRA of 67 who claims at age 62 receives only 70% of their FRA benefit amount.

A worker may choose to defer receipt of SSB past his FRA. In this case a delayed retirement credit (DRC) will be added to the FRA benefit. A worker with a FRA of 66 who delays claiming until age 70 receives 132% of their FRA benefit amount; a worker with a FRA of 67 who claims at age 70 receives only 124% of their FRA benefit amount.

Workers who claim early retirement benefits, but continue to work, may have their SSB reduced. This is referred to as the Earnings Test (ET). However, since 2000, there has been no ET above the FRA.[1] That is, SSB are not reduced if the worker is of FRA and continues to work.

A spouse has dual entitlements to SSBs. A spouse is entitled to the larger of 100% of benefits at FRA based on his or her earnings record or up to 50% of the spouse's benefits at FRA.

$$SSB_{spouse1} = \text{Max}\ \{SSB_{own};\ .5(SSB_{spouse2})\}$$

[1] http://www.socialsecurity.gov/pubs/10003.html

Once one begins SSBs based on his or her own work record they cannot later switch to SSBs based on the spouse's record. Also, one cannot begin SSB based on the spouse's record and then later switch to SSBs based on his or her own work record. However, there is an exception: a wife (husband) can retire and begin collecting her (his) own SSBs while her (his) husband (wife) still works and delays benefits. Upon her (his) husband's (wife's) retirement, she (he) can switch over to 50% of his (her) benefits, if spousal benefits are greater than her (his) own benefits. Spouse's benefits do not include any accrued delayed retirement credits.

For example, assume Richard and Jane, are both 62 with a FRA of 66. Currently, Richard's SSB at FRA are $2,000 per month and Jane's SSB at FRA are $1,000. Jane retires at 62 and receives 75% of 1,000 or $750 per month. Richard continues to work until age 66. His SSB at FRA are still $2,000 per month and he retires at FRA. Assuming no COLA for Jane's SSB, she can now switch over to spousal benefits of 50% x $2,000 = $1,000 per month.

MODEL

Similar to McCormack and Perdue (2006), we avoid the problem of an uncertain discount rate (DR) by computing the internal rate of return (IRR) equating two retirement options. For married couples of the same age, the IRR can be solved for by using the following equation:

$$\%\text{Benefit\_1} \times \sum_{1}^{i} \left(\frac{1}{1 + \frac{\text{IRR}}{12}} \right)^{i} + \%\text{Benefit\_2} \times \sum_{1}^{j} \left(\frac{1}{1 + \frac{\text{IRR}}{12}} \right)^{j}$$

$$= \%\text{Benefit\_3} \times \sum_{1}^{m} \left(\frac{1}{1 + \frac{\text{IRR}}{12}} \right)^{m} \times \left(\frac{1}{1 + \frac{\text{IRR}}{12}} \right)^{N3-N1}$$

$$+ \%\text{Benefit\_4} \times \sum_{1}^{n} \left(\frac{1}{1 + \frac{\text{IRR}}{12}} \right)^{n} \times \left(\frac{1}{1 + \frac{\text{IRR}}{12}} \right)^{N4-N2}$$

where:

Benefit_X = percent of SSB received based on retirement age
i = 1 to months to life expectancy for retirement Age 1 male (N1)
j = 1 to months to life expectancy for retirement Age 1 female (N2)
m = 1 to months to life expectancy for retirement Age 2 male (N3)
n = 1 to months to life expectancy for retirement Age 2 female (N4)
N3 – N1 and *N4 – N2* = difference in months between retirement Age 1 and retirement Age 2, where retirement Age 2 is greater than retirement Age 1.

The two terms on the left-hand side of the equation,

$$\%\text{Benefit\_1} \times \sum_{1}^{i} \left(\frac{1}{1 + \frac{\text{IRR}}{12}} \right)^{i} \quad \text{and} \quad \%\text{Benefit\_2} \times \sum_{1}^{j} \left(\frac{1}{1 + \frac{\text{IRR}}{12}} \right)^{j},$$

represent the present value of initiating receipt of benefits at retirement age 1. The two terms on the right-hand side of the equation,

$$\%\text{Benefit\_3} \times \sum_{1}^{m} \left(\frac{1}{1 + \frac{\text{IRR}}{12}} \right)^{m} \quad \text{and} \quad \%\text{Benefit\_4} \times \sum_{1}^{n} \left(\frac{1}{1 + \frac{\text{IRR}}{12}} \right)^{n},$$

represent the present value of initiating receipt of benefits at retirement age 2; the two second terms on the right-hand side,

$$\left(\frac{1}{1 + \frac{\text{IRR}}{12}} \right)^{N3-N1} \quad \text{and} \quad \left(\frac{1}{1 + \frac{\text{IRR}}{12}} \right)^{N4-N2},$$

discount the present value of benefits at retirement age 2 back to retirement age 1 so that the IRR can be computed at the same point in time. For example, if the first retirement age is 62 and the second retirement age is 66, the IRR computation for the age 66 term must be discounted back to the same point in time as the age 62 term.

It should be noted that this model is appropriate only for same aged couples retiring at the same age. When the couples are different ages but still retire at the same age, an additional discount factor $\left(\frac{1}{1+\frac{\text{IRR}}{12}} \right)^{D}$ is required to discount all expected cash flows back to the initial start of benefits. The model now becomes:

$$\%\text{Benefit\_1} \times \sum_{1}^{i} \left(\frac{1}{1 + \frac{\text{IRR}}{12}} \right)^{i} + \%\text{Benefit\_2} \times \sum_{1}^{j} \left(\frac{1}{1 + \frac{\text{IRR}}{12}} \right)^{j} \times \left(\frac{1}{1 + \frac{\text{IRR}}{12}} \right)^{D}$$

$$= \%\text{Benefit\_3} \times \sum_{1}^{m} \left(\frac{1}{1 + \frac{\text{IRR}}{12}} \right)^{m} \times \left(\frac{1}{1 + \frac{\text{IRR}}{12}} \right)^{N3-N1}$$

$$+ \%\text{Benefit\_4} \times \sum_{1}^{n} \left(\frac{1}{1 + \frac{\text{IRR}}{12}} \right)^{n} \times \left(\frac{1}{1 + \frac{\text{IRR}}{12}} \right)^{N4-N2} \times \left(\frac{1}{1 + \frac{\text{IRR}}{12}} \right)^{D}$$

where:

D = the age difference in months between the spouses ($Age_{husband} - Age_{wife}$) and $Age_{husband} > Age_{wife}$.

In addition, if the couples are different ages and retire at different ages, additional discounting complications are introduced. The model now becomes:

$$\%\text{Benefit\_1} \times \sum_{1}^{i}\left(\frac{1}{1+\frac{\text{IRR}}{12}}\right)^{i} + \%\text{Benefit\_2} \times \sum_{1}^{j}\left(\frac{1}{1+\frac{\text{IRR}}{12}}\right)^{j} \times \left(\frac{1}{1+\frac{\text{IRR}}{12}}\right)^{D-(N1-N2)}$$

$$= \%\text{Benefit\_3} \times \sum_{1}^{m}\left(\frac{1}{1+\frac{\text{IRR}}{12}}\right)^{m} \times \left(\frac{1}{1+\frac{\text{IRR}}{12}}\right)^{N3-N1}$$

$$+ \%\text{Benefit\_4} \times \sum_{1}^{n}\left(\frac{1}{1+\frac{\text{IRR}}{12}}\right)^{n} \times \left(\frac{1}{1+\frac{\text{IRR}}{12}}\right)^{N4-N2} \times \left(\frac{1}{1+\frac{\text{IRR}}{12}}\right)^{D-(N3-N4)}$$

4.1. Assumptions in the Model

The following assumptions are made:

1) SSB are received monthly.
2) The retirement decision is made annually because life expectancy tables only provide annual data.
3) The 2006 United States Life Tables and the 2010 National Center for Health Statistics provide life expectancies.[2]

[2] National Vital Statistics Report, June 28, 2010, Volume 58, Number 21; United States Life Tables, 2006 provides life expectancies for black and white males and females. Arias E., United States life tables by Hispanic origin. National Center for Health Statistics. Vital Health Stat 2(152). 2010 provides life expectancies for Hispanic males and females.

Table 1: Average life expectancy given current age

Age	White Males Avg # years remaining	Black Males Avg # years remaining	Hispanic Males Avg # years remaining
62	19.32	16.90	21.26
63	18.57	16.29	20.48
64	17.83	15.69	19.71
65	17.10	15.10	18.96
66	16.38	14.51	18.21
67	15.67	13.93	17.48
68	14.97	13.36	16.77
69	14.28	12.80	16.07
70	13.60	12.25	15.38

Age	White Females Avg # years remaining	Black Females Avg # years remaining	Hispanic Females Avg # years remaining
62	22.18	20.72	24.24
63	21.37	19.99	23.39
64	20.56	19.27	22.55
65	19.76	18.57	21.72
66	18.97	17.87	20.90
67	18.18	17.17	20.10
68	17.41	16.48	19.30
69	16.64	15.80	18.51
70	15.89	15.14	17.74

Source: National Vital Statistics Report, June 28, 2010, Volume 58, Number 21; United States Life Tables, 2006; and Arias E. United States life tables by Hispanic origin. National Center for Health Statistics. Vital Health Stat 2(152). 2010.

Life expectancy is adjusted for when a worker retires. For example, a white male who retires at age 62 is expected to live approximately 19 more years to age 81; whereas if he waits and retires at age 66 he is expected to live approximately 16 more years to age 82. We look at life expectancies based on gender and race.

4) We assume excess earnings are $0 and that early retirement SSB are not further reduced by the earnings test.
5) If a retiree has substantial income (earned and unearned) in addition to his SSB, up to 85% of his annual benefits may be subject to Federal income tax. In our analysis we assume other income is below the minimum such that 0% of SSB are taxed. However, by using the IRR method to find the optimal retirement age, taxation of SSB really becomes irrelevant, since (1-tax rate of SSB) shows up on both the left- and right-hand sides of our equation, effectively cancelling out one another.
6) Since 1983, the SSA provides for an automatic increase in SSB if there is an increase in the CPI-W from third quarter last year to third quarter of the current year. Spitzer (2006) finds that only longevity and expected rates of return are determining factors as the optimal time to retire and that inflation and taxes play no significant role. As a consequence, we assume COLA is zero.
7) We assume the husband (male) is older than the wife (female). We look at age differences ($Age_{male} - Age_{female}$) of 0, 4, 7, and 10. This assumption will be relaxed in future studies, allowing the wife to be older than the husband.
8) We assume a one-earner family. The husband is the working spouse, and the wife is the non-working spouse. Thus, a wife receives one-half of her husband's full retirement benefit unless the wife begins collecting benefits before her FRA. If the wife begins collecting benefits before her FRA, the amount of the wife's benefit is reduced by a percentage base on the number of months before she reaches FRA. For example, based on the FRA of 66, if the wife begins collecting benefits:
   At age 65, the benefit amount would be about 45.8 percent of the retired worker's (husband's) full benefit;
   At age 64, it would be about 41.7 percent;
   At age 63, 37.5 percent; and
   At age 62, 35 percent.
   This assumption will be relaxed in future research to allow the wife to be the working spouse, and allow a two-earner family.
9) We also assume the couple has no dependents, and that neither party receives a government pension. Furthermore, the couple may be forced into a higher federal or state tax bracket due to other income; this, too, is irrelevant in our analysis and is ignored.

AN EXAMPLE

Let us look at Michael, a black male born in 1952, who is trying to decide if he should retire early at age 62 or wait until his FRA of 66. Michael is married to Angela, a black female

born in 1952, who has no SSB of her own. According to Table 1, Michael's life expectancy at age 62 is an additional 16.90 years (202.8 months) to age 78.9; while his life expectancy at age 66 is an additional 14.51 years (174.12 months) to age 80.51. Angela's life expectancy at age 62 is an additional 20.72 years (248.64 months) to age 82.72; while her life expectancy at age 66 is an additional 17.87 years (214.44months) to age 83.87. Based on current Social Security requirements, Michael will receive 100% of his SSB at age 66, but only 75% of his FRA benefits at age 62. Angela is able to claim up to 50% of Michael's SSB if she is at FRA, but only 35% at age 62.

Using Excel and Solver we can find the IRR that will equate both sides of the following equation:

$$75\% \times \sum_{1}^{202.8} \left(\frac{1}{1 + \frac{\text{IRR}}{12}} \right)^{i} + 35\% \times \sum_{1}^{248.64} \left(\frac{1}{1 + \frac{\text{IRR}}{12}} \right)^{j}$$

$$= 100\% \times \sum_{1}^{174.12} \left(\frac{1}{1 + \frac{\text{IRR}}{12}} \right)^{m} \times \left(\frac{1}{1 + \frac{\text{IRR}}{12}} \right)^{(66-62)\times 12}$$

$$+ 50\% \times \sum_{1}^{214.44} \left(\frac{1}{1 + \frac{\text{IRR}}{12}} \right)^{n} \times \left(\frac{1}{1 + \frac{\text{IRR}}{12}} \right)^{(66-62)\times 12}$$

The IRR that equates both sides is equal to 5.53%. If the couple's opportunity costs are less (greater) than 5.53%, then they should retire at the later (earlier) age.

Assume Michael's SSB at FRA of 66 is $1,600 per month and his early retirement benefit is 75% or $1,200 per month at age 62. Based on Michael's FRA benefit of $1,600 per month, Angela's SSB will be 35% of $1,600 or $560 per month at age 62. At age 66 Michael will receive $1,600 per month and Angela will receive 50% of $1,600 or $800 per month. If the current market interest rate is 5%, then the present value (PV) of the left-hand side of the equation (retire early at age 62) is $164,070 (Michael) plus $86,603 (Angela) for a total of $250,673. The PV of the right-hand side of the equation (delay retirement to age 66) is $162,038 (Michael) and $92,787 (Angela) for a total of $254,825. This results in a difference of $4,152, implying that that Michael and Angela should wait until age 66 to retire. If Michael and Angela believe they could invest their monthly SSB at 5.53% or greater over the next four years, then they should retire early, at age 62; if not, they should delay retirement until age 66. Of course, this assumes they do not need any of their SSB on which to live - a highly unlikely assumption.

In this example, with no differences in age and retiring at the same age, the breakeven IRR is 5.53% (5.5291% from Table 2 rounded to 2 decimal places). However, if the wife is 4 years

younger there is an additional 4 years of discounting required (48 months) for the wife spousal benefits at both age 62 and 66. This is reflected in the following formula:

$$75\% \times \sum_{1}^{202.8} \left(\frac{1}{1+\frac{\text{IRR}}{12}}\right)^{i} + 35\% \times \sum_{1}^{248.64} \left(\frac{1}{1+\frac{\text{IRR}}{12}}\right)^{j} \times \left(\frac{1}{1+\frac{\text{IRR}}{12}}\right)^{(4\times 12)}$$

$$= 100\% \times \sum_{1}^{174.12} \left(\frac{1}{1+\frac{\text{IRR}}{12}}\right)^{m} \times \left(\frac{1}{1+\frac{\text{IRR}}{12}}\right)^{(66-62)\times 12}$$

$$+ 50\% \times \sum_{1}^{214.44} \left(\frac{1}{1+\frac{\text{IRR}}{12}}\right)^{n} \times \left(\frac{1}{1+\frac{\text{IRR}}{12}}\right)^{(66-62)\times 12} \times \left(\frac{1}{1+\frac{\text{IRR}}{12}}\right)^{(4\times 12)}$$

Note that the age 62 spousal benefits are now discounted 48 months (instead of none previously) and the age 66 spousal benefits are now discounted 96 months instead of 48 months. Using Excel and goal seek we find the breakeven IRR is 5.4042% which is reflected in Table 2 with a 4 year age difference.

To illustrate an example from Table 3 again consider the same couple above with a 4 year age difference but with the H/W early retirement ages of 66/62 and delayed retirement ages of 70/66. The formula to solve this example would be:

$$100\% \times \sum_{1}^{174.12} \left(\frac{1}{1+\frac{\text{IRR}}{12}}\right)^{i} + 35\% \times \sum_{1}^{248.64} \left(\frac{1}{1+\frac{\text{IRR}}{12}}\right)^{j} \times \left[\left(\frac{1}{1+\frac{\text{IRR}}{12}}\right)^{[4-(66-62)]\times 12}\right]$$

$$= 132\% \times \sum_{1}^{147} \left(\frac{1}{1+\frac{\text{IRR}}{12}}\right)^{m} \times \left(\frac{1}{1+\frac{\text{IRR}}{12}}\right)^{(70-66)\times 12}$$

$$+ 50\% \times \sum_{1}^{214.44} \left(\frac{1}{1+\frac{\text{IRR}}{12}}\right)^{n} \times \left(\frac{1}{1+\frac{\text{IRR}}{12}}\right)^{(66-62)\times 12} \times \left[\left(\frac{1}{1+\frac{\text{IRR}}{12}}\right)^{[4-(70-66)]\times 12}\right]$$

Using goal seek to solve for the breakeven IRR yields 4.7984% (see Table 3).

RESULTS

Table 1 provides the average life expectancies for both males and females for the three race categories (White, Black and Hispanic) that the Breakeven (BE) Internal Rates of Return (IRR) in Tables 2 and 3 are based on. The results presented in Tables 2 and 3 are based on applying the previously described Excel model for a representative baby boom birth year of 1948 for both the husband and wife initially and progressively later years for the non-working female spouse.

Table 2 provides the BE IRR's for the 9 race combinations where W = White, B = Black and H = Hispanic. The 9 husband/wife combinations are: WW, BB, HH, WB, BW, WH, HW, BH, HB. Following Blanchett (2013), three claiming scenarios are considered: receiving benefits early (e.g., at age 62 versus 66); the maximum realistic delay period (e.g., at age 62 versus 70) and delaying benefits past full retirement age (e.g., age 66 versus 70). Within these 3 claiming scenarios we examine couples by race combination who retire at the same age with age differences of 0, 4, 7 and 10 years with the non-working spouse younger than the assumed working husband. These assumptions are, admittedly, arbitrary but useful from our perspective to examine the impact of increasing age differences on the breakeven IRR's.

Table 2 Breakeven IRRs for a Sample of Married Retirement Ages With Increasing Age Differences

Age Difference	Male Retirement Age1	Female Retirement Age1	Male Retirement Age2	Female Retirement Age2	WB Breakeven IRR	BW Breakeven IRR	WH Breakeven IRR
0	62	62	66	66	5.4566%	5.5371%	5.6061%
4	62	62	66	66	5.3301%	5.4115%	5.4601%
7	62	62	66	66	5.2168%	5.2977%	5.3317%
10	62	62	66	66	5.0862%	5.1661%	5.1851%
0	62	62	70	70	2.9863%	3.0080%	2.9830%
4	62	62	70	70	3.0656%	3.1025%	3.0615%
7	62	62	70	70	3.1848%	3.2385%	3.1864%
10	62	62	70	70	3.4207%	3.5005%	3.4391%
0	66	66	70	70	0.3148%	0.1922%	-0.0123%
4	66	66	70	70	0.3446%	0.2137%	-0.0137%
7	66	66	70	70	0.5873%	0.4828%	0.2525%
10	66	66	70	70	1.2727%	1.2629%	1.0592%

Table 2 Breakeven IRRs for a Sample of Married Retirement Ages With Increasing Age Differences

Age Difference	Male Retirement Age1	Female Retirement Age1	Male Retirement Age2	Female Retirement Age2	HW Breakeven IRR	BH Breakeven IRR	HB Breakeven IRR
0	62	62	66	66	5.6842%	5.6786%	5.6773%
4	62	62	66	66	5.5730%	5.5342%	5.5667%
7	62	62	66	66	5.4717%	5.4065%	5.4666%
10	62	62	66	66	5.3554%	5.2606%	5.3520%
0	62	62	70	70	3.3166%	3.0957%	3.3998%
4	62	62	70	70	3.4230%	3.1836%	3.5000%
7	62	62	70	70	3.5607%	3.3166%	3.6294%
10	62	62	70	70	3.8089%	3.5789%	3.8630%
0	66	66	70	70	0.7710%	0.1275%	0.9983%
4	66	66	70	70	0.8504%	0.1429%	1.0912%
7	66	66	70	70	1.1270%	0.4297%	1.3586%
10	66	66	70	70	1.8220%	1.2686%	2.0017%

Keep in mind that the BE IRR's can be viewed as "hurdle rates" where if a couple's expected return or opportunity cost of capital is greater than (less than) the computed BE IRR over the given time horizon, the couple should retire at the earlier (later) age. This analysis also assumes that the couple does not need the social security benefits to live on and can invest the benefits in the capital markets if the decision is made to retire early.

Our results are surprisingly similar across the 9 race combinations, but are different for the 3 age group comparisons. For the age 62 versus 66 comparisons the BE IRR's uniformly decrease as the age difference increases. Since, as noted above, these IRR's are hurdle rates, this implies that greater age difference couples should retire earlier since the hurdle rate is less to overcome than at a smaller age difference. These results should be interpreted with caution however since an inflection points occurs at the age 62 versus 67 comparison (not shown) and continues onto the age 62 versus 70 comparison where the IRR's uniformly increase with age differences. We attribute this inflection point to the interaction of an increasing time gap between the early and delayed retirement with a constant set of age differentials. This age 62 versus 70 comparison implies that greater age differences involve a greater hurdle and the smaller the age difference the greater the incentive to retire earlier since the hurdle rate is lower. The results for the age 66 versus 70 comparison are similar to the age 62 to 70 comparison with the breakeven IRR's increasing with age differences although the numbers themselves are quite small by comparison and would seem to suggest early retirement at all age differences given the low hurdle rates to overcome.

It is also interesting and useful to compare the results across race categories at key comparison ages. From Table 2, the high and low breakeven Internal Rates of Return for the following retirement age comparisons are evident:

Retirement Age Comparison	High Breakeven IRR	Low Breakeven IRR
62/62 versus 66/66		
Age Difference		
0	HH	WB
4	HH	WB
7	HH	WB
10	HH	WB
62/62 versus 70/70		
Age Difference		
0	HB	WW
4	HB	WW
7	HB	WW
10	HH	WW
66/66 versus 70/70		
Age Difference		
0	HB	WH
4	HB	WH
7	HB	WH
10	HB	WW

Recall that W = White, B = Black and H = Hispanic and the husband is listed first and the wife second. So, for example, WB refers to a white husband married to a black spouse. Note that a higher (lower) Breakeven IRR would imply retiring later (earlier) since the hurdle rate opportunity cost is more difficult (less difficult) to overcome. The high breakeven IRR column is dominated by HB (7 occurrences) and HH (5 occurrences). The low breakeven IRR column has 5 WW lows, 4 WB lows and 3 WH lows. The most obvious patterns here are the fact that the High Breakeven IRR group consistently has a Hispanic husband and Low Breakeven IRR group consistently has a white husband. For a given retirement age comparison/age difference the results can be interpreted as follows: the high (low) breakeven group would prefer to retire later (earlier) since the hurdle rate is more difficult (less difficult) to overcome.

In Table 3 we examine breakeven IRR's for couples by race combination who retire at different ages (Table 2 assumed the same retirement age for the couples) and who have a positive age difference. There is a Not Applicable (NA) in the table for an age difference of 0 since spousal benefits cannot be claimed by the female until the male retires. Table 3 examines the impact of age differences on an early male/female retirement of 66 and 62 respectively versus a late male/female retirement of 70 and 66 respectively. In all 9 race combinations the breakeven IRR's decline as the age differences increase. This suggests that the greater the age difference the greater the incentive to retire early as the hurdle rate is lower to overcome. It is also interesting to examine the high and low breakeven IRR's for this comparison for each age difference by race category:

Retirement Age Comparison 66/62 versus 70/66 Age Difference	High Breakeven IRR	Low Breakeven IRR
4	HH	WB
7	HH	WB
10	HH	WB

Interestingly, in all cases the HH (WB) race combination has the highest (lowest) breakeven IRR's. This suggests a later (earlier) retirement for the HH (WB) race combinations when comparing for a given age difference since the hurdle rate is higher (lower) respectively for the 2 combinations.

Table 3 Breakeven IRRs for a Sample of Married Retirement Ages With Different Retirement Ages and Increasing Age Differences

Age Difference	Male Retirement Age1	Female Retirement Age1	Male Retirement Age2	Female Retirement Age2	WW Breakeven IRR	BB Breakeven IRR	HH Breakeven IRR
0	66	62	70	66	NA	NA	NA
4	66	62	70	66	4.5980%	4.7984%	5.0782%
7	66	62	70	66	4.4554%	4.6613%	4.9375%
10	66	62	70	66	4.2894%	4.5026%	4.7772%

Table 3 Breakeven IRRs for a Sample of Married Retirement Ages With Different Retirement Ages and Increasing Age Differences

Age Difference	Male Retirement Age1	Female Retirement Age1	Male Retirement Age2	Female Retirement Age2	WB Breakeven IRR	BW Breakeven IRR	WH Breakeven IRR
0	66	62	70	66	NA	NA	NA
4	66	62	70	66	4.5863%	4.8090%	4.7454%
7	66	62	70	66	4.4453%	4.6704%	4.5878%
10	66	62	70	66	4.2816%	4.5096%	4.4063%

Table 3 Breakeven IRRs for a Sample of Married Retirement Ages With Different Retirement Ages and Increasing Age Differences

Age Difference	Male Retirement Age1	Female Retirement Age1	Male Retirement Age2	Female Retirement Age2	HW Breakeven IRR	BH Breakeven IRR	HB Breakeven IRR
0	66	62	70	66	NA	NA	NA
4	66	62	70	66	4.9467%	4.9523%	4.9375%
7	66	62	70	66	4.8207%	4.7987%	4.8129%
10	66	62	70	66	4.6756%	4.6224%	4.6698%

APPLICATIONS/IMPLICATIONS

The practical applications/implications of our results primarily depend on the couple's opportunity cost of capital and available other resources. If the couple's portfolio expected return or opportunity cost of capital is greater than (less than) the computed breakeven IRR, this would suggest that this couple retire at the earlier (later) date in the comparative analysis. These results should be useful for couples of different ages facing the Social Security early versus delayed retirement decision and financial planners. Using the analytics described in this paper, couples

and/or their financial planners could first compute their breakeven Internal Rates of Return at various comparison ages and then compare this breakeven IRR to their expected portfolio return over the comparison period. If their expected portfolio return was greater than (less than) their breakeven IRR then they should consider retiring at the earlier (later) age.

CONCLUSIONS

The primary substantive conclusions from this study depends on the age comparisons that are being made. For different aged couples who retire at the same chronological age, the age 62 versus 66 comparisons show BE IRR's uniformly decrease as the age difference increases. Since these IRR's are hurdle rates, this implies that greater age difference couples should retire earlier since the hurdle rate is less to overcome than at a smaller age difference. These results reverse for the age 62 versus 70 comparison and age 66 versus 70 comparisons where the IRR's uniformly increase with age differences across all race combinations. This implies that greater age differences involve a greater hurdle and the smaller the age difference the greater the incentive to retire earlier since the hurdle rate is lower. For couples who have an early male/female retirement of 66 and 62 respectively versus a late male/female retirement of 70 and 66 respectively the breakeven IRR's consistently decline as the age differences increase across all race combinations. This suggests that the greater the age difference the greater the incentive to retire early as the hurdle rate is lower to overcome.

REFERENCES

Arias E. United States life tables by Hispanic origin. National Center for Health Statistics. Vital Health Stat 2(152). 2010.

Blanchett, David. (2012). When to Claim Social Security Benefits. *Journal of Personal Finance* 11:2, 36-87.

Coile, Courtney, Diamond, Peter, Gruber, Jonathan, and Jousten, Alain. (2002). Delays in Claiming Social Security Benefits. *Journal of Public Economics*, 84, 357-385.

Docking, Diane Scott., Fortin, Richard, and Michelson, Stuart. (2012). The Influence of Gender and Race on the Social Security Early Retirement Decision for Single Individuals. *Journal of Economics and Economic Education Research*, 13:1, 87-104.

Docking, Diane Scott., Fortin, Richard, and Michelson, Stuart. (2013). The Influence of Race on the Social Security Early Retirement Decision for Married Couples. *Financial Services Review* 22:2, 133-150.

McCormack, Joseph P. and Perdue, Grady. (2006). Optimizing the Initiation of Social Security Benefits. *Financial Services Review*, 15, 335-348.

Munnell, Alicia H. and Soto, Mauricio. (2007). When Should Women Claim Social Security Benefits? *Journal of Financial Planning*, 20:6, 58-65.

National Vital Statistics Report, June 28, 2010, Volume 58, Number 21; United States Life Tables, 2006

Social Security Administration. Available at http://www.socialsecurity.gov/pubs/10003.html

Spitzer, John J. (2006). Delaying Social Security Payments: A Bootstrap. *Financial Services Review*, 15, 233-245.

Sun, Wei and Webb, Anthony. (2009). How Much Do Households Really Lose by Claiming Social Security at Age 62? *Center for Retirement Research at Boston College*, CRR WP 2009-11, March, 1-28.

Tucker, Michael. (2009). Optimal Retirement Age Under Normal and Negative Market Conditions Considering Social Security and Private Savings. *Journal of Financial Planning*, July. 42-49.

U.S. Census Bureau. Available at http://www.census.gov/population/www/socdemo/age/general-age.html#bb.

APPENDIX A

Abbreviation	Meaning
COLA	Cost of Living Adjustment
DR	Discount Rate
DRC	Delayed Retirement Credit
ERA	Early Retirement Age
ET	Earnings Test
FRA	Full Retirement Age (receive full 100% of benefits)
IRR	Internal Rate of Return
PV	Present Value
SSA	Social Security Administration
SSB	Social Security Benefit

THE ALCHIAN-ALLEN EFFECT IN HIGHER EDUCATION REVISITED: STATE LOTTERY IMPACTS ON PUBLIC VERSES PRIVATE ENROLLMENT[1]

James B. Larriviere[2], Spring Hill College
Anindya Biswas,Spring Hill College

Abstract: The purpose of this paper is to revisit the Alchian-Allen effect in higher education and to examine the impact of state lotteries on the public/private enrollment ratio. The Alchian-Allen effect applies to situations where there is a change in consumption resulting from the change in relative prices due to a fixed cost the consumer must pay regardless of the which product is purchased. Kroncke and Ressler (1993) tested for the Alchian-Allen effect in higher education and concluded the price ratio of public to private education increased with a decrease in the fixed cost of attending higher education, proxied by the unemployment rate, bringing about a decrease in the public to private enrollment ratio. Our results also indicate the unemployment rate does have a significant direct effect on the enrollment ratio (confirming the Alchian-Allen effect); however, the existence of a state lottery tends to decrease the public/private enrollment ratio and therefore reduces impact of the Alchian-Allen effect in higher education enrollment.

JEL classifications: D12, I23

Keywords: Alchian Allen (AA) Effect, Public versus private enrollment in higher education, State lotteries.

1. INTRODUCTION

The Alchian Allen (AA) Effect, coined by Alchian and Allen (1968), has been applied to a plethora of situations where the relative prices of two substitute goods change as a result of a fixed cost that the buyer must pay. The classic example of the AA effect was posited by Borcherding and Silberberg (1978). The question considered was why the better quality apples were shipped out of Washington, leaving the lower quality apples to be consumed within the state. Borcherding and Silberberg answer the question by explaining the existence of shipping costs, a fixed cost, that when applied to the apples, will make the more expensive apple relatively cheaper, and therefore more attractive. Consequently, the better quality apples are a better bargain to the out of state consumers than to in-state consumers.

Kroncke and Ressler (1993) tested the existence of the AA effect in higher education, examining two goods - public and private higher education. Using the unemployment rate as a

1 The authors would like to thank Rand Ressler and Starla Paccini for helpful comments on earlier drafts of the paper, but take responsibility for any remaining errors.

2 Corresponding Author

proxy for the opportunity cost of attending post-secondary education, Kroncke and Ressler conclude the relative price of public and private education does change as a result of changes in foregone employment opportunities. Changes in unemployment rates impact the fixed cost, and therefore the full cost, of attending both public and private institutions, hence changing the relative price between the two. When the relative prices change, so will the quantities demanded of each good.

One rather recent development that has impacted higher education is the existence of state lotteries. Currently, forty three states have state lotteries, and fifty percent of these lotteries began after 1986. Lotteries have been marketed by their supporters as the panacea for the lack of education funding provided by the state. With state budgets under constant pressure, lottery tax revenue is a welcome supplement to the state's spending on education, potentially impacting the relative prices of public and private education.

The purpose of this paper is to examine the AA effect that lotteries may have on the U.S. higher education. First, whether the AA effect of unemployment rate still exists in higher education is examined. The model specification developed by Kroncke and Ressler (1993) is the basis for the model utilized in this paper. Second, a new model specification will be estimated utilizing a new variable to control for the existence of state lotteries. By accounting for the existence of lotteries, we examine whether the existence of state lotteries enhances the AA effect, or whether it in fact reduces the AA effect on the public and private enrollment ratio.

2. THE MODEL AND FINDINGS

A basic cross-sectional ordinary least square (OLS) model is utilized to determine if the AA effect exists in higher education. The "fixed cost" workers experience is measured by the state's unemployment rate, which proxies the opportunity cost of attending college. As the unemployment rate increases, the probability of getting a job decreases, and thus the opportunity cost (forgone income) of enrolling in higher education falls. If the AA effect does exist, the unemployment rate will have a positive sign, indicating as the opportunity cost of going to college decreases (higher unemployment rate), then public enrollment (the assumed lower quality good) should increase to a greater degree than the private school enrollment.

The OLS specification of this basic model is:

$$Enrollratio = \gamma + \beta 1 Unemplrate + \beta 2 Income + \beta 3 Finaid + \beta 4 Pop + \beta 5 Tuitionratio$$

where *Enrollratio* = the ratio of public enrollment to private enrollment of each state, γ is the intercept, *Unemplrate* = the unemployment rate of each state, *Income* = the per-capita income of each state, *Finaid* = financial aid distributions of each state, *Pop* = the 18-24 year old population of each state, *Tuitionratio* = state's ratio of public institution tuition to private institution tuition. This variable measures the Department of Education's obligations for student financial assistance. The data consists of an observation for each state for the fiscal years 2004 and 2005. The variables are in their natural logarithmic form, except for the unemployment rate and the tuition ratio.

Table 1: All models

Regressors	Model 1 Coefficient estimate (p-value)	Model 2 Coefficient estimate (p-value)	Model 3 Coefficient estimate (p-value)
γ	21.708 (0.000)	21.499 (0.000)	18.355 (0.001)
Unemplrate	18.336 (0.017)	17.914 (0.011)	15.356 (0.024)
Income	-1.782 (0.001)	-1.779 (0.001)	-1.479 (0.003)
Finaid	-0.139 (0.091)	-0.138 (0.013)	-0.101 (0.062)
Pop	-0.003 (0.980)	-	-
Tuitionratio	-0.426 (0.643)	-	-
Lottery	-	-	-0.628 (0.004)
F	6.470 (0.000)	10.070 (0.000)	11.190 (0.000)
B-P χ^2	10.070 (0.073)	8.170 (0.043)	6.730 (0.151)
p-value of S-W W test	(0.383)	(0.279)	(0.074)
adj-R^2	0.246	0.262	0.327

Note. This table provides the results of this analysis for the three models: Model 1, Model 2, and Model 3. In all models we try to explain average enrollment ratio of public versus private education. The results show that Model 3 explains better the variation in enrollment ratio and it does not suffer from the heteroskedasticity problem.

The second column, labeled as Model 1, of the Table 1 contains the results of this basic model. The existence of the AA effect is evident with a statistically significant, positive coefficient on *Unemplrate*. As the unemployment rate decreases, the ratio of public to private enrollment decreases as well. The $Income$ coefficient is negative and highly significant indicating that as state per-capita income increases, the enrollment ratio declines, either by more students enrolling in private institutions, fewer students enrolling in public institutions, or possibly the increase in private sector enrollment increasing to a greater degree than the public sector enrollment. Another income-related variable, *Finaid*, has a significant negative coefficient. The effect of financial aid assistance seems to have a stronger impact on enrollments in private institutions relative to public institutions. The *Tuitionratio* and *Pop* coefficients are negative, although insignificant. This model specification supports the notion that the unemployment rate has an AA effect on the enrollment ratio in higher education, reaffirming the finding of Kroncke and Ressler (1993).

Given the fact our data are cross-sectional, it is necessary to test for the occurrence of heteroskedasticity in the residuals of the model. The well-known Breusch-Pagan (B-P) test for heteroskedasticity is utilized and the p-value of the B-P χ^2 statistic is reported on the 9th row of Table 1. The results from this test indicate we are unable to reject the null hypothesis of constant variance as the corresponding p-value is 0.07 (> 0.05). Since we are assessing significance of five independent variables by corresponding *t*-tests we also need to check the normality assumption of the residuals of the model. For this purpose the Shapiro-Wilk (S-W) W test is used to test for normality, and the corresponding p-value is 0.38 (> 0.05) indicating we cannot reject the null hypothesis the residual is normally distributed. Thus, it appears Model 1 does not violate any of the standard assumptions of the OLS. However, we cannot be certain without a further analysis of the adequateness of Model 1 specification for explaining the Alchian-Allen effect in higher education enrollment.

A second model is estimated that omits the two insignificant variables in the first model specification: *Tuitionratio* and *Pop*.

$$Enrollratio = \gamma + \beta 1 Unemplrate + \beta 2 Income + \beta 3 Finaid$$

The results of this model are presented in the third column, labeled as Model 2, in Table 1. The adjusted R-squared improves by 6.5% with two fewer explanatory variables, and all three of the explanatory variables are significant at the 5% level. The unemployment rate variable is once again positive, showing the existence of the AA Effect. The Income variable is once again negative, indicating as income increases, consumers of higher education choose to enroll in private institutions to a greater degree than public institutions. The financial aid variable is also inversely related to the enrollment ratio. As with the income variable, it appears the financial assistance has a greater impact on private enrollment relative to public enrollment.

An examination and testing of common OLS assumptions of Model 2 does identify some concerns. The p-value of the B-P χ^2 statistic, reported on the 9th row, third column of Table 1, indicates that the underlying model may suffer from the heteroskedasticity problem, as the corresponding p-value is 0.04 (< 0.05). However, the corresponding p-value of the S-W W test is 0.28 (> 0.05) indicating we cannot reject the null hypothesis that the residual is normally distributed. We also note that t-tests for the three independent variables in the model are highly significant as p-values are less than 0.05 and the VIF's (< 10) are small (see Table 2) indicating insignificant multicollinearity. Therefore, Model 2 does not violate the normality assumption, but may suffer from the heteroskedasticity problem.

Given the heteroskedasticity concern, the residuals of Model 1 and Model 2 specifications are examined to detect if there is any important independent variables omitted from those models. Basic econometric theory states there should not be any pattern apparent in the residuals plot, meaning econometrically meaningful models' residuals should bear only a random pattern. Initially, these residual plots (see left hand side diagram of first two panels of Figure 1) seem to exhibit a rough random pattern, implying that no adjustment to these models can be made to improve these regression equations. However, after careful consideration we noticed an interesting and fairly consistent pattern in the residuals plot of these models: states that do not have a lottery (represented by the dummy variable and is equal to zero) are positive

with just one exception (see right hand side of the first two panels of Figure 1). This pattern implies the OLS models under-predicted the enrollment ratio. Thus, it appears we can improve the fit of these models by adding a variable that represents the lottery effect.

2.1. STATE LOTTERY IMPACT ON HIGHER EDUCATION ENROLLMENT

Our results so far have confirmed the existence of the Alchian - Allen Effect in higher education; the unemployment rate in the state has a significant, positive impact on the public/private enrollment ratio. As the unemployment rate declines, the enrollment ratio decreases, meaning private enrollment is increasing to a greater degree than public enrollment. The results from our analysis of the residuals from Model 1 and Model 2 indicate there are strong econometric reasons, in addition to mere economic intuition, to include the lottery variable in our model.

Currently, lotteries are in forty-three states, with Alabama, Alaska, Hawaii, Mississippi, Nevada, Utah, and Wyoming currently not operating state lotteries. Among the forty-three states operating a lottery the appropriation of lottery tax revenue varies. Some states allocate the revenue to the general fund, whereas other states earmark the revenues to public K-12 education, or programs for the elderly, as well as scholarship programs in higher education. For a majority of the states, the revenues are generally split between the state's common fund and public K-12 education. Only a small percentage of states with lotteries explicitly earmark lottery revenue to benefit higher education. Florida and Georgia, among others, use a portion of the lottery tax revenue to fund various merit-based and need-based scholarship programs (Borg and Borg, 2007). Researchers have explored the impact of lottery proceeds on the incidence of the benefits of lottery-funded HOPE scholarships that Georgia residents received (Rubenstein and Scafidi 2002). Dyarski (2000) examined the impact of lottery-funded merit-based scholarships on in-state college attendance rates of middle and upper income students. We contribute to this literature by specifying a model (Model 3) that includes a dummy variable for the lottery. The question we examine here is whether the existence of a lottery in a state has an impact on the public/private enrollment ratio. We also examine the impact the lottery has on the AA effect. We include a dummy variable, *lottery*, which has a value of 1 if the state has a state lottery and 0 if no lottery exists in the state. The lottery variable does not account for how the lottery revenue is apportioned; it simply accounts for the existence of a lottery in the state. The Model 3 is specified as follows:

$$Enrollratio = \gamma + \beta 1 Unemplrate + \beta 2 Income + \beta 3 Finaid + \beta 4 lottery$$

The results of the third model specification are presented in the fourth column, labeled as Model 3, of Table 1. The increase in the adjusted R-square by 25% relative to Model 2 and by 33% relative to Model 1 supports our choice of including the lottery variable in Model 3. The same tests for heteroskedasticity and normality were used in Model 3, and the results indicate that we failed to reject both the null hypotheses of homoscedasticity and normality of the residuals distribution. The lottery variable is significant at the 1% level, and is negatively related to the enrollment ratio. Moreover, we do not find any consistent pattern in the residuals plot of Model 3 (see diagrams of third panel in Figure 1). The results show the existence of a lottery in the state decreases the public to private enrollment ratio. One possibility is the lottery increases private enrollment to a greater degree than public enrollment. It is also possible that public

enrollment is declining as the private enrollment is increasing. This result may support the view the public education is an inferior good compared to the private education. In explaining this view, it could be the existence of state lotteries reduce the overall tuition cost the student faces, leading to an increase in effective income. This increase in available income leads to more private education being purchased than public education.

The findings by Rubenstein and Scafidi (2002) may also help us to explain the inverse relationship between the lottery variable and the public/private enrollment ratio. Rubenstein and Scafidi concluded that white, higher-income households receive a disproportionately large number of lottery funded, merit-based scholarships Therefore, it is possible states with lotteries would see more scholarships awarded to higher income households, and these households could be choosing the more selective, "higher quality" private institutions.

Including the lottery variable in Model 3 does affect the size of the Alchian-Allen effect relative to Model 1 and Model 2. The smaller coefficient on the unemployment rate indicates the size of the AA Effect is slightly reduced when the lottery is considered. The coefficients on income and the financial aid variable both decreased to some degree with the inclusion of the lottery variable. It appears the effect the lottery has on the enrollment ratio accounts for a small portion of both the income effects and financial aid effects on the enrollment ratio.

4. CONCLUSION

This paper has tested for and confirmed the existence of the Alchian-Allen effect in higher education. Our results show as the opportunity cost of attending college (proxy by the unemployment rate) changes, the relative full prices of public to private education change, thereby changing the quantity demanded of both goods. The second contribution of this paper is the inclusion and examination of a lottery variable and its impact on the public to private enrollment ratio. The results from the model including the lottery variable confirm that in addition to the unemployment rate having a significant direct effect on the enrollment ratio, the existence of a state lottery has a statistically significant, negative impact on the public to private enrollment ratio, indicating the lottery is associated with a decrease in the public to private enrollment ratio. Specifically, the private enrollment increases to a greater degree than public enrollment in states that have a lottery. Further study on this lottery effect on higher education enrollment could attempt to explain how lottery revenue appropriation plans impact the public to private enrollment ratio.

REFERENCES

Alchian, A. and Allen, R. (1968) University Economics, second edition. Belmont: Wadsworth Publishing Company.

Borg, J.R. and Borg, Mary. The Reverse Robin Hood Effect: The distribution of Net Benefits from the Florida Bright Futures Scholarship. Florida Political Chronicle 1, Fall 2007.

Burcherding, T.E. and Silberberg, E. (1978) Shipping the Good Apples Out: The Alchian and Allen Theorem Reconsidered. Journal of Political Economy 86, 131-138.

Dynarski, S. (2000) Hope for Whom? Financial Aid for the Middle Class and its Impact on College Attendance. National Tax Journal 53, 629-661.

Kroncke, C. 0. and Ressler, R. W (1993). The Alchian-Allen Effect in Higher Education: Public Versus Private Enrollment. Economics of Education Review, 12, 345-349.

National Association of State Student Grant and Aid Programs. 37th Annual Survey Report on State Sponsored Student Financial Aid. 2005-2006 Academic Year.

Rubenstein, Ross and Scafidi, Benjamin. Who Pays and Who Benefits? Examining the Distributional Consequences of the Georgia Lottery for Education. National Tax Journal 55, 223-238.

United States Department of commerce, Bureau of the Census. Statistical Abstract of the United States. Washington DC: U.S. Government Printing Office, various years.

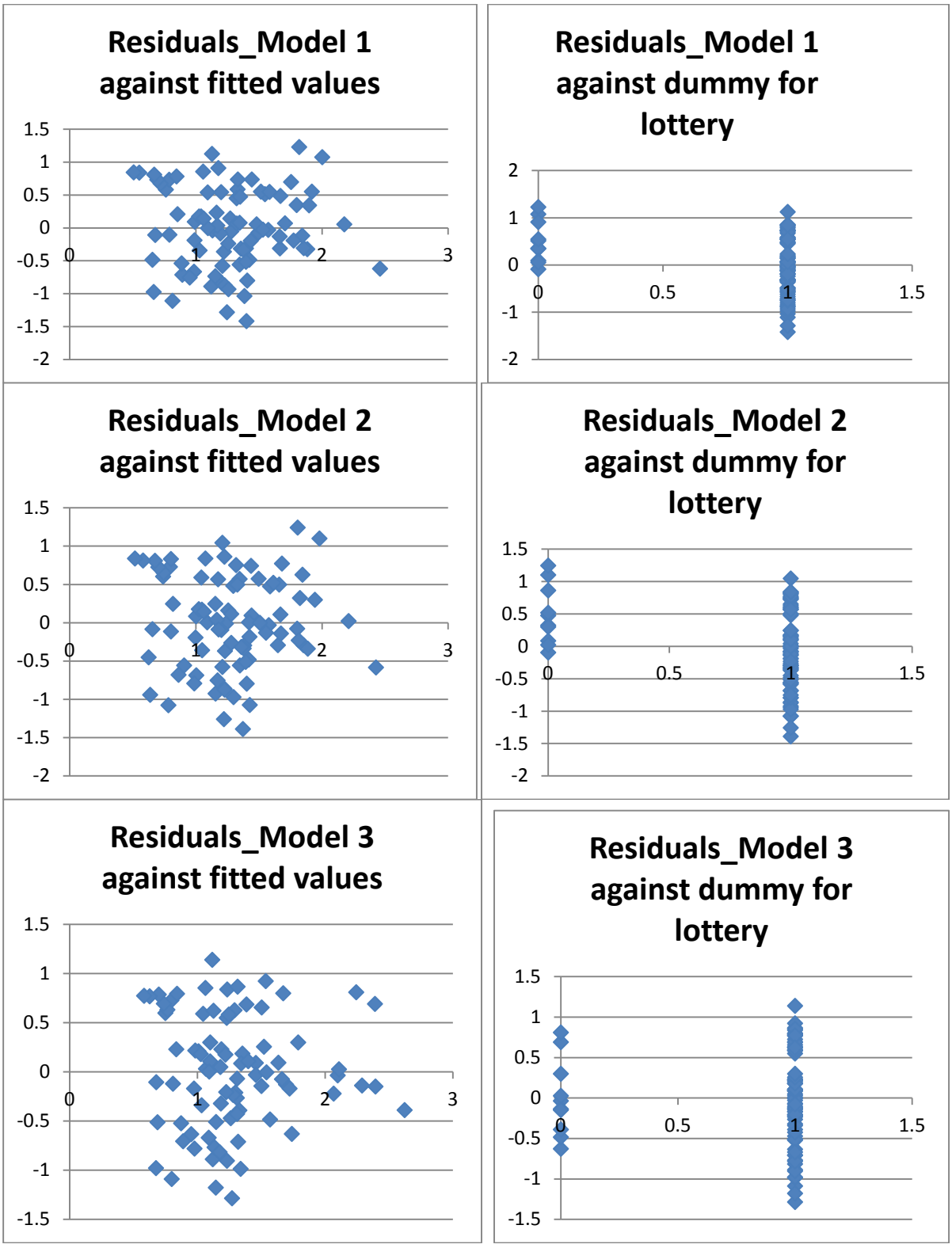
U.S. Department of Education: Institute of Education Services, National Center for Education Statistics. Digest of Education Statistics 2007, 2008.

Table 2: Variance Inflation Factor

Variables	Model 1 VIF	Model 2 VIF	Model 3 VIF
Unemplrate	1.31	1.13	1.15
Income	1.17	1.10	1.15
Finaid	2.32	1.06	1.12
Pop	2.73	-	-
Tuitionratio	1.06	-	-
Lottery	-	-	1.14

Note: This table provides the variance inflation factor of all the explanatory variables for the three models: Model 1, Model 2, and Model 3. The results show that none of the models suffer from the multicolinearity problem.

Figure 1


Residuals_Model 1 against fitted values
Residuals_Model 1 against dummy for lottery
Residuals_Model 2 against fitted values
Residuals_Model 2 against dummy for lottery
Residuals_Model 3 against fitted values
Residuals_Model 3 against dummy for lottery

RESIDUALS PLOT

This figure plots the residuals of the three models: Model 1, Model 2, and Model 3 on the left hand side of the three panels against fitted values. On the right hand side it plots residuals against the dummy variable for lottery. The diagrams on the right hand side of the first two panels show a fairly consistent and interesting pattern: states those have no lotteries are positive with just one exception whereas the third panel shows that residuals for states with no lotteries have random positive and negative values.

EFFECT OF INVENTORY MANAGEMENT EFFICIENCY ON PROFITABILITY: CURRENT EVIDENCE FROM THE U.S. MANUFACTURING INDUSTRY

Seungjae Shin, Mississippi State University
Kevin L. Ennis, Mississippi State University
W. Paul Spurlin, Mississippi State University

ABSTRACT

While manufacturing firms pursue efficient inventory management, there is limited evidence of improved financial performance related to inventory management practices. This paper examines financial statement data for U.S. manufacturing firms to explore the relationship between inventory management efficiency and firm profitability. The results show that a lower ratio of inventory to sales for a firm is associated with higher profit margin for the firm. In addition, small size firms can receive a larger benefit (as measured by profitability) from increased inventory efficiency when compared to medium and large size firms.

Key words: Inventory Management, Profitability, U.S. Manufacturing Industry

INTRODUCTION

Maintaining an appropriate level of inventory is a key issue to firms' operational performance. The supposition is that better inventory management is closely related with firms' better financial performance. Appropriate inventory levels depend on the production schedule as a managerial response to market demand. Inventory is a current asset to a firm, but it is costly to maintain as it waits to be converted into future sales. While excess inventory does increase costs, a shortage of inventory may result in lost sales. Prior research has focused on inventory management methods and optimal inventory sizes as they relate to the balance between more technological information systems, inventory cost savings and production/sales efficiency. Inventory management has evolved into a highly studied and practiced concept in the business world that combines optimizing inventory movement, information-sharing between buyer and seller, lean production strategies, and supply chain management concepts. The core of the current inventory management system is Just-In-Time (JIT) inventory systems.

JIT is a philosophy of management that reduces waste and improves quality in all business process (Harrison and Hoek, 2011). JIT has been applied to many Japanese manufacturing firms since the 1970s (Cheng and Podolsky, 1996). JIT originated from the Toyota production system (TPS) and serves to reduce inventory and lead-time while increasing quality of production. JIT is defined as, "an inventory strategy aimed at improving a business' financial performance by reducing excess inventory and its associated cost" (Sungard, 2007). To implement a JIT inventory system, a sound, long-term relationship with suppliers is critical because suppliers have to fill the inventory as soon as it reaches a minimum level. Therefore, sharing information about the production schedule with part suppliers and delivery companies is

essential. This information sharing is now available through a modern IT infrastructure utilizing the Internet and Enterprise Resource Planning (ERP).

ERP was introduced in the 1990s as an enterprise information system designed to integrate production and accounting data and functions across organizations. The main goal of ERP is to share data by all functional departments and to access the data immediately to increase prompt decision making (Motiwalla and Thompson, 2009). Together, the Internet and ERP systems dramatically improve the JIT inventory system, allowing real time information tracking and sharing of both production and accounting information. JIT inventory management and the utilization of Internet and ERP systems provides for a "lean production" opportunity.

The concept of lean production is to minimize inventory and has been widely used since the 1990s (Eroglu and Hofer, 2011). JIT is the heart of the lean production systems. In the late 1990's, the JIT and ERP concepts expanded into a concept known as Supply Chain Management (SCM). The supply chain is defined as "management of network of interconnected business," to satisfy customers' requests (Harland, 1996).

As stated above, the implementation of a technological complete inventory management system to determine an appropriate or optimal inventory level is a critical factor to a firms' financial performance. Better inventory management such as higher inventory turnovers, reduced days-in-inventory, or lower level of inventory-to-sales ratio is closely related with firms' better financial performance (Shah & Shin, 2007). Using data collected from the late 1960's; the late 1990's; and some early 2000's, prior studies investigated the relationship between inventory level and firm's financial performance. A sample of these studies is delineated in our next section of this paper. This prior research offers both numerous and conflicting results as both positive relationships and negative relationships were determined. In addition, inventory management and its impact on financial performance based on firm size was not considered. Because a definitive answer does not exist as to whether optimizing inventory management is related to superior firm financial performance and does the impact differ based on manufacturing firm size, this paper investigates whether successfully managing a low level of inventory will result in higher profitability for the firm. These conflicting relationships and lack of information of the impact on firm size coupled with our utilization of more recent data from U.S. manufacturing firms leads us to a single research hypothesis which states:

Hypothesis: A significant relation exists between firm profitability and inventory management efficiency in U.S. manufacturing industry.

This study begins with a brief literature review followed by data collection methods, research analysis, and a conclusion.

LITERATURE REVIEW

The concept of Supply Chain Management and technologically managing inventory has helped a lot of companies to compete more effectively in their business markets. Kannan and Tan (2004) point out the three popular methods used in order to ensure that the product or service is delivered to the customer in the most efficient way possible. These three methods are JIT, Total Quality Management (TQM), and SCM. All three of these methods go hand in hand because they force the company to eliminate waste while increasing the quality of their products and distribution systems. Their research demonstrates that integrated inventory management

methods are correlated with firm financial performance. Using return of assets (ROA) as a measure of financial performance, Kannan and Tan (2004) set out to not only reiterate the impact on firm's operational performance, but also point out that the firm's business performance can benefit from an inventory management system. Their results concluded that integrating a technological inventory management system results in higher ROA.

Shah and Shin (2007) investigate the relationship among IT investment, inventory, and financial performance with industry sector level data of 1960 to 1999. They find that lower inventory levels lead to higher financial performance in manufacturing sector. Their conclusion is that there exists indirect effect on financial performance through inventory management from IT investment. Liberman and Demeester (1999) study Japanese car manufacturers' JIT production with data of late 1960s to early 1990s. They find that there is a causal relationship between work-in-process inventory and firm's productivity, i.e., 10% reduction in inventory leads to 1% increase of labor productivity. Thomas (2002) studies inventory changes and future returns with data from 1970 through 1997. He finds that a firm with inventory increase has experiences higher level of profitability, however, this trend changes immediately with a change of inventory decrease. He finds the negative relationship between inventory level and firm's profitability but he cannot explain the reason. The result from Thomas (2002) conflicts with results from both Liberman and Demeester (1999) and Shah and Shin (2007). Chen, Frank and Wu (2005) investigate inventories of U.S. manufacturing companies in the last two decades of 20th century. They find that firms with high inventory have poor long-term stock returns while firms with slightly lower than average inventory have good stock returns. However, firms with lowest inventory have only normal returns. All four papers study about the relationship with financial performance of U.S. manufacturing industry. But their results are not consistent. The data used in the previous four papers are data of 20th century. Roumiantsev and Netessine (2007) investigated linkage of inventory behavior with financial performance. They found that lower inventory levels are positively associated with return on sales. Capkun et al. (2009) found a significant positive correlation between inventory performance and measures of financial performance in manufacturing companies over 26 year period from 1980 to 2005.

Profitability is a concept that a lot of executives and shareholders put emphasis on. This shows them that their company is operating at a level to where more money is coming in than leaving the company. Gill, Biger, and Muthur (2010) discusses the relationship that occurs between the firm's working capital management and profitability. They define working capital as being involved with current assets and current liabilities while being able to finance these current assets. The main difference between inventory management and working capital management is the fact that working capital management involves managing all of the current assets while inventory management focuses its efforts on inventories alone. Gill et al. (2010) stated that they do not see any relationship between days of accounts payable and profitability or even with days in inventory and profitability. They note that past studies have given results that differ from their own. Given conflicting results in previous studies, we are motivated to offer evidence as to whether inventory management and profitability are related.

DATA

Using the *Compustat* database, the authors obtain annual balance sheet and income statement data for US manufacturing firms. Manufacturing firms are identified in the database by using the NAICS (North America Industry Classification System) code. Manufacturing

companies have a NAICS code beginning with 31, 32, and 33. The authors conduct tests on two sets of data: The first data set is made of three years of data from 2005 and 2007, and the second expands the time window to eight years, from 2005 to 2012. Table 1 presents the two data sets.

Table 1
Data Sets

	Data Set I	Data Set II
Periods	2005 ~ 2007 (3 years)	2005 ~ 2012 (8 years)
Number of Firms	1,289	959
Number of Observation	3,867	7,672

The total number of manufacturing companies, which are listed in the U.S. stock markets in 2005 is 1,292. Among them, 1,289 U.S. manufacturing firms are listed for three years from 2005 and 959 firms are listed for eight years. The number of firms decreases in our sample when the sample period is increased to eight years because we use a balanced panel and annual observations for sample firms that are available for all years in the first sample are not available for all firms in the second sample. Data set I ends with fiscal year 2007 in an attempt to avoid our data being influenced by the financial crisis beginning in 2008. During the financial crisis 2008 through 2009, many companies were delisted (Erkens, Hung, and Matos, 2012). To form a balanced panel of data, we require observations for sample firms to occur in all three or eight years and to include all financial statement items so that our final sample includes a total of 1,289 / 959 firms times 3 years / 8 years. Profit margin (PM), calculated as the ratio of net income to total revenue, is used as our measurement of a company's profitability, and inventory-sales-ratio (ISR), calculated as total inventory divided by total revenue, is used to measure inventory management efficiency with a lower ISR being interpreted as a higher level of inventory management efficiency.

ANALYSIS I

This study uses cross-sectional, time-series panel data. Cross-sectional variables are ISR and PM collected for 1,289 U.S. manufacturing companies. The time series for these variables is collected for three consecutive years. The total number of observations is 3,867. We classified these data into three groups based on size of year-firm revenue (table 2): (1) Small size companies with less than or equal to $100 million dollars per year, (2) medium size company with an annual revenue between $100 million dollars and $1 billion dollars, and (3) large size company with greater than or equal to $1 billion dollars per year. Using this convention, a firm may appear as a different size from one year to the next in our sample. The authors assumed that company's inventory management efficiency depends on the size of the company. The average ISR for the large companies is 0.1265 and those for the medium and small companies are 0.154 and 0.229 respectively.

Table 2
Classification of Company Size and average ISR

Size	Criteria	Number of Observations	Average ISR
Small	≤ $ 100 M	1030	0.228890
Medium	Between $100 M and $ 1 B	1554	0.154085
Large	≥ $1B	1283	0.126500

Estimating the panel data regression model by ordinary least square (OLS) might provide a biased solution caused by unobserved heterogeneity (Dougherty, 2006). To overcome this possible problem, two approaches were offered: fixed effect and random effect. According to Green (2012), while the fixed effect assumes that individual heterogeneity is correlated with independent variables, the random effect assumes that the individual heterogeneity is uncorrelated with the independent variables. Jerry A. Hausman developed a test for determining which model is appropriate.

To estimate the panel data regression model in this study, the following procedure will be used: (1) Estimate the regression equation with the assumption that intercepts and slope coefficients are constant across time and individual companies, which is called the pooled regression model. (2) Do the Hausman Test to find a better approach between fixed effect model and random effect model. (3) Based on the result in the (2), explore the better model than the pooled regression model.

As we mentioned earlier, we choose ISR as a variable for inventory management efficiency and PM as a variable for company's profitability. In addition, we introduced four dummy variables: two intercept dummies and two slope dummies. The reason for dummy variables is that each size of company may have a different effect on the profitability by inventory management efficiency. To test this relation, the authors estimated the following models:

$$PM_{i,t} = \beta_0 + \beta_1 Large_{i,t} + \beta_2 Medium_{i,t} + \beta_3 ISR_{i,t} + \beta_4 Large_{i,t}*ISR_{i,t} + \beta_5 Medium_{i,t}*ISR_{i,t} + u_{i,t} \quad (1)$$

Where

$Large_{i,t}$ = 1 if the company i belongs to large size company at year t, 0 otherwise

$Medium_{i,t}$ = 1 if the company i belongs to medium size company at year t,0 otherwise

The table 3 presents output from the above regression model by OLS. The F-statistic is 204.7 with a p-value of 2.2e-16. The R^2 is 0.2095.

Table 3
Result from Regression by OLS

	β_i	t-value	p-value	α
Intercept	3.5918	6.296	3.40e-10	0.001
Large	-3.5142	-3.009	0.002637	0.001
Medium	-3.5338	-3.365	0.000774	0.001
ISR	-28.1524	-31.649	< 2e-16	0.001
Large ISR*	28.0762	3.9159	20e-05	0.001
Medium ISR*	27.9826	5.529	2.44e-08	0.001

The regression model (1) needs to be tested for the assumption of homeskedasticity and no serial correlation. To test the homoskedasticity assumption, the Breusch-Pagan (BP) Test was used. The BP Test is supposed to detect heteroskedasticity by running a regression with the squared residuals as a dependent variable. Because the *p-value* (9.98e-11) was less than .01, the null hypothesis of homoscedasticity was rejected with a 1% significance level. Therefore, this data was heteroskedastic. The problem with heteroskedasticity is that the t-statistics of coefficients cannot be trusted because the estimated standard errors are biased. The heteroskedasticity-consistent (HC) standard errors procedure was proposed by Halbert White to fit a model with heteroskedastic residual. This correction procedure for HC standard errors is

called White correction. The output from the HC standard errors has the same coefficient with different *t-values* and *p-values* of the coefficients, which are listed in table 4.

Table 4
Result from Regression by HC Standard Errors

	β_i	t-value	p-value	α
Intercept	3.5918	5.059583	0.0e+00	0.001
Large	-3.5142	-4.950047	1.0e-06	0.001
Medium	-3.5338	-4.977455	1.0e-06	0.001
ISR	-28.1524	-4.270610	2.0e-05	0.001
Large ISR*	28.0762	4.258905	2.1e-05	0.001
*Medium * ISR*	27.9826	4.244723	2.2e-05	0.001

To test the autocorrelation, the Durbin-Watson (DW) test was used. The Durbin-Watson (DW) statistic (*d*) is 1.8968 with a p-value of 0.0005753. The DW *d* statistic must be compared to two critical d values: d_L and d_H. With six independent variables and 2,000 observations, the 1% one-tailed critical values are d_L = 1.89104 and d_H = 1.90106. If the *d* is greater than d_H, the null hypothesis, no evidence of positive correlation, cannot be rejected. Variance Inflation Factor (VIF) is a measure of multicollinearity, which is caused by highly correlated independent variables. The suggested cutoff value for VIF is 5.0. The VIF values of five variables (table 5) were between 1.0 and 3.5, i.e. no multicollinearity.

Table 5
Variance Inflation Factor

Large	Medium	ISR	Large*ISR	Medium*ISR
3.996999	3.504987	1.064515	3.427672	2.964264

Result from the Hausman test is 10.1009 chi-square statistic with a p-value of 0.07243, which is greater than the significance level 5%. The Hausman test tells that the null hypothesis of no preference between two models cannot be rejected. Therefore, the random effect model was used. The table 6 presents the output from the random effect model. The F Statistic is 1591.38 with a p-value of 2.22e-16 and the R^2 is 0.67332.

Table 6
Result from Random Effect Model

	β_i	t-value	p-value	α
Intercept	3.41011	6.296	3.40e-10	0.001
Large	-3.68276	-3.009	0.002637	0.001
Medium	-3.26944	-3.365	0.000774	0.01
ISR	-23.9909	-31.649	< 2e-16	0.001
Large ISR*	20.55188	3.9159	20e-05	0.001
*Medium * ISR*	15.70121	5.529	2.44e-08	0.001

Table 7 presents values of intercept and slope for each category. Since values of slope are negative, PM of all three categories rise when ISR goes down. The small size company has the highest intercept value and lowest slope value, which means that the small size company has a

greatest effect of ISR on PM among the three categories, indicating all three cases supports the *Hypothesis.*

Table 7 **Intercept and Slope for Each Category**		
	Intercept	Slope
Small Size Company	3.41011 (β_0)	-23.9909 (β_3)
Medium Size Company	0.14067 ($\beta_0+\beta_1$)	-8.28989 ($\beta_3+\beta_4$)
Large Size Company	-0.27265 ($\beta_0+\beta_2$)	-3.43902 ($\beta_3+\beta_5$)

ANALYSIS II

The purpose of analysis II is to add robustness to our result from analysis I by using an expanded time period for our data set and by using a firm's level of assets as a different measure of firm size in order to control for firm profitability that may be related to firm size. To test the relation between profitability and inventory management efficiency, we estimate the following cross-sectional, time series, balanced panel model:

$$PM_{i,t} = \beta_0 + \beta_1 LN(Assets_{i,t}) + \beta_2 ISR_{i,t} + u_{i,t} \quad (2)$$

According to Eriotis, Frangouli, and Ventoura-Neokosmides (2002), we include natural log value of total assets, $LN(Assets_{i,t})$, as a control variable which is firm *i*'s natural log value of total assets in year *t*. The total number of observations in the data set was 7,672, which is data for 959 companies for 8 years from 2005 through 2012. We use the same procedure as we did in the analysis I. The Breusch-Pagan (BP) test is used to test the homoskedasticity and DW test is used to check for autocorrelation. In addition, a Hausman test is used to choose between random effect model and fixed effect model. Because the *p-value* from the BP test was near zero (<.0001), the null hypothesis of homoskedasticity is rejected with a 1% significance level. Therefore, this data qualified as heteroskedastic. The Durbin-Watson (DW) statistic, *d*, is quite low, 0.901. With four independent variables including an intercept term and 2000 observations, the 1% one-tailed critical values were $d_L = 1.89405$ and $d_H = 1.89804$. Because $d < d\ L$, our data likely exhibits autocorrelation. Because of the near zero *p-value (< .0001)* from the Hausman test, the null hypothesis is rejected. Therefore, the fixed effect model is more appropriate. Table 8 presents fixed effect regression results.

Table 8 **Result from Fixed Effect Model**			
	Coefficient	T-Value	p-Value
Intercept	-2.67865	-5.59	< .0001
LN(Assets)	0.139846	6.72	< .0001
ISR	-4.76086	-111/05	< .0001

The result of analysis II shows that a significantly negative relation exists between a firm's ISR and the firm's PM. A smaller value for the ISR ratio indicates the firm is more efficient at managing inventory. Therefore, the negative relation between firm PM and ISR supports the *Hypothesis* as we find that firms with better inventory management efficiency show

higher profitability. The R^2 from our regression estimation is 0.7495, indicating that our model has a strong fit. The F-statistic of 7.99 for our model is significant at the 0.01% level. The result from the analysis II validates the result from analysis I, i.e., there is a negative relationship between ISR and PM. Therefore, our results indicate that efficient inventory management practices have a positive effect on firm's profitability.

CONCLUSION

While inventory management has been common practice in business for many years, there has been an evolution for inventory management from JIT to lean inventory systems to supply chain management. However, the supposition about the cost/benefits of the implementation of inventory management is both conflicting and dated. The purpose of this study is to determine if a relationship exists between inventory management efficiency and firm profitability. First, to be more precise, the logical analysis is that the management of the appropriate inventory levels would result in an inverse relationship between the optimized inventory levels and firm profitability. Second, we attempt to expand the investigation to determine the impact of inventory efficiency by firm size and, finally, offer corroborative evidence by using more recent data. Our results show a positive relation between profitability and inventory management efficiency. In addition, the impact of that inventory efficiency on profitability based on firm size was noticeably significant.

In the first model, we control for firms size with binary dummy variables concerning revenue. In the second model, we control for firm size with a continuous variable measuring asset size. Both models show same result for the relation between ISR and profitability. While both models significantly support the inverse relationship between lower inventory levels and increase in profitability, the results for the firm size are significantly noticeable. In the first model, the result shows that the smaller category of firm size has a stronger negative effect on its profitability. However, in the second model (based on asset size) and overall general comparison of manufacturing firms, there is a positive effect of firm size on profitability. This adds corroborating evidence to the overall analysis and supportive evidence to the inventory level/profitability relationship. From our results that are robust to time period and measurement variation, we conclude that more efficient inventory management practices result in higher firm profitability.

Finally, it is important to include further analysis concerning the impact of inventory efficiency and firm size. As seen in Table 2, the impact of inventory efficiency is potential greater in small size firms as opposed to medium and large size firms. In other words, it is possible to conclude that small size firms can receive a larger benefit (as measured by profitability) from increased inventory efficiency when compared to medium and large size firms. It could be concluded that the potential for improvement is greater in the smaller firms. One possible explanation for this is that the medium and larger manufacturing firms have already implemented fully technological advanced inventory management systems and have already maximized its potential effect of inventory efficiency. The smaller firms could have greater room for improvement by increasing inventory efficiency.

REFERENCES

Capkun, V., A. Hameri, and L.A. Weiss, (2009). On the relationship between inventory and financial performance in manufacturing companies, *International Journal of Operations & Production Management*, 29 (8), 789-806.

Cheng, T.C.E. and S. Podolsky. (1996). *Just-In-Time Manufacturing*, 2nd e., New York, NY: Chapman and Hall USA.

Dougherty, C. (2006). *Introduction to Econometric 3rd Edition*, New York: Oxford University Press.

Eriotis, N. P., Z. Frangouli, and Z. Ventoura-Neokosmides (2002). Profit Margin and Capital Structure: An Empirical Realationship *Journal of Applied Business Research, 18(2), 85-88.*

Erkens. D.H., M. Hung, and P. Matos. (2012). Corporate Governance in the 2007-2008 Financial Crisis: Evidence from Financial Institutions Worldwide, *Journal of Corporate Finance*, 18 (2) 389-411.

Eroglu, C. and C. Hofer. (2011). Lean, leaner, too lean? The inventory-performace link revisited, *Journal of Operations Management*, 29, 356-369.

Gill, A., Biger, N., and Mathur, N. (2010). The Relationship between Working Capital Management and Profitability: Evidence from the United States. *Business and Economics Journal*, 1-9.

Green, W.H. (2012). *Econometric Analysis 7th e.*, Upper Saddle River, NJ: Prentice Hall.

Harland, C.M. (1996) Supply Chain Management, Purchasing and Supply Management, Logistics, Vertical Integration, Materials Management and Supply Chain Dynamics. In: Slack, N (ed.) *Blackwell Encyclopedic Dictionary of Operations Management*. UK: Blackwell.

Harrison, A. and Hoek, R. V. (2011). *Logistics Management and Strategy: Competing Through the Supply Chain, 4th e*., Harlow, England: Prentice Hall.

Liberman, M. and L. Demeester. (1999). Inventory Reduction and Productivity Growth: Linkages in the Japanese Automotive Industry, *Management Science*, 45 (4), 466-85.

Kannan, V. R. and K. H. Tan. (2005), Just in Time, total quality management, and supply chain management: understanding their linkages and impact on business performance. *Omega*, 33 (2), 153-162.

Motiwalla, L. and J. Thompson, (2009), *Enterprise Systems for Management*, Upper Saddle River, New Jersey: Pearson Education, Inc.

Roumiantsev, S. and S. Netessine, (2007), Inventory and its relationship with Profitability: Evidence for an International Sample of Countries: Working Paper, Retrieved Feb. 1, 2014 from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2319862

Sungard, (2007), Information Resiliency for a JIT world, Retrieved Feb. 1, 2014, from http://www.informationavailability.com/NR/rdonlyres/CBA20DDC-332C-42DF-95D9-693729993E20/0/JITWP.pdf

Shah, R. and H. Shin. (2007). Relationships among information technology, inventory and profitability: An investigation of level invariance using sector level data, *Journal of Operation Management*, 15, pp 768-84.

Thomas, J. (2002). Inventory Changes and Future Returns, *Review of Accounting Studies*, 7, 163-187.

White, H. (1980). A heteroskedasticity-consistent covariance matrix and a direct test for hetero-skedasticity, *Econometrica*, 48 (4): 817-838.

WAGE DIFFERENTIALS FOR PRE-COLLEGE TEACHERS TO SOLVE SHORTAGES AND QUALITY ISSUES?

Paul M. Mason, University of North Florida
Michael Fabritius, Centre College
Christopher K. Johnson, University of North Florida

ABSTRACT

Complaints about the quality of schools abound, but solutions are few and far between. This note presents what we believe is one comprehensive solution to most of the ills of public and private schools, namely wage differentials for teachers that reflect their respective opportunity costs of being educators. The analysis reveals graphically, and with historical data, how both teaching and non-teaching markets are distorted by salary scales and why this leads to disproportionate numbers and qualities of teaching faculty across disciplines. The end of the paper develops justifications for how and why this will improve schools.

JEL Classification: I21, I22

Key Words: wage differentials, high opportunity cost teachers, low opportunity cost teachers, Finnish schools

INTRODUCTION

The fundamental premise of this communication is that many of the basic problems inherent in secondary education could be alleviated through wage differentials for teachers. Unfortunately, there have historically been no wage differentials in most public educational institutions at the primary or secondary level and the private middle and high schools that do have wage differentials protect their structure and design with stringent confidentiality requirements. In addition, it would be virtually impossible to distinguish the wage effect from all others that fuel the perception that private schools are better than public ones (or vice-versa).

In the last few years, discussion of merit pay for teachers and differential pay based on student performance has surfaced, but not the pervasive wage differentials discussed here. Consequently, if general wage differentials are ever applied at the public, pre-college level, they will have to be developed from scratch for that segment of the educational structure. Naturally, wage differentials are not a goal, but rather a tool to achieve what is most desirable, that is a climate in which student learning is maximized. This can be accomplished when classroom teachers are the best available and the atmosphere for the development of the highest order analytical and cognitive skills is possible in each student to his/her ability to achieve.

A quote in *The Economist* (March 25-31, 2006, p. 58) reveals the simple explanation from a headmistress of a school near Helsinki, Finland for why the Finnish schools were identified as the world's best (see Alvarez, 2004 referenced in *Freakonomics* 2006 for a discussion of the Finnish superiority). She

says "[T]eachers, teachers, teachers." Finnish schools have either eliminated, or avoided instituting national curriculums, magnet schools, charter schools, national exams, streaming, or selection programs. Instead, they provide considerable support to teachers and grant teachers the freedom to operate most effectively. It appears to be working. However, maintaining quality in the classroom requires proper incentives universally.

BRIEF LITERATURE REVIEW

Up until very recently, there has been very little literature regarding the role of wage differentials in primary and secondary educational institutions. Except for a small group of journal articles in the late 1970s, there have been only a few analyses in the popular press and from think tanks regarding the standard salary structure found in most of our public and private schools today. An article by the Pacific Research Institute questioned why teacher salaries are not tied to either market demand or classroom performance. "An English teacher makes the same salary as a physics teacher. Although there is a lot of discussion about a general teacher shortage, the real scarcity is in specific fields such as math and science. Individuals with college degrees in math and science can get much higher salaries in private industry than they can get in teaching." (Izumi, 2001) A large part of the rationale for this reality is the inability of policy makers and educators to think outside the box of uniform wage rates for all teachers. This severely hampers any effective effort to address this shortage (Izumi, 2001).

Despite claims that teaching is a calling; and that teachers most value the intrinsic rewards, teachers react similarly to other workers in terms of compensation. Teachers respond to wage differentials between teaching and non-teaching jobs (Baugh and Stone, 1982). Thus, teachers' salaries represent an important policy issue that contributes to a school's ability to attract qualified staff. Baugh and Stone (2011) recently expanded this discussion in their book based on the 1982 article.

A study conducted by Dale Ballou (2000) relative to teacher contracts in Massachusetts, found widespread agreement that teacher salaries were not sufficiently competitive to attract talented people into careers in education in the numbers that are desired. He found this especially true of teachers who possess knowledge and skills that are in high demand in private industry. Administrators that were contacted for Ballou's study frequently commented on how difficult it has become to recruit teachers in mathematics and science given the premium on their skills in today's labor market, yet the structure of teacher compensation in Massachusetts has remained resistant to market forces (Ballou, 2000). Consequently, regardless of what subject they teach, teachers are paid according to the same set of criteria: essentially how many advanced degrees or college credits they have beyond a particular degree level, and how many years of service. There is generally little in the way of merit pay, no differentials by field, no differentiation to compensate teachers for taking on jobs in particularly difficult working conditions (Ballou, 2000).

Another study completed by the National Center for Policy Analysis (1999) found, like the Finns, that the single most important factor in student learning improvement is teacher quality. Yet teacher quality is woefully missing in certain high school subject areas. There is significant evidence that teacher quality varies by class level, with the advanced classes being allocated the superior teachers (see Ingersol, 1999). The research also shows that teacher knowledge of specific subject matter, particularly at the secondary level, is a good prediction of student achievement. For example, David Monk (1994) found a strong correlation between teacher preparation in both math and physical sciences and student success in both low and high scoring students.

Within the last twenty months, two additional analyses have been published that indirectly assess wage differentials for teachers by evaluating the by-discipline earnings of those teachers that have left the profession. Chingos and West (2012) employ a unique data set of 3,500 ex-teachers from the state of Florida. The results suggest that math and science teachers who leave teaching earn respectively 15% and 11.8% more than departed English teachers in alternative employment. West (2013) sites his prior work and that of a few others to infer that compensating high demand teachers more might keep them from departing and attract other qualified teachers from alternative employment. However, to be impartial,

West also cites evidence from several sources revealing the reluctance of parents to support wage differentials as opposed to higher pay for all teachers, the potential for a perspective of lesser importance for other disciplines as a result, and issues associated with union contracts if such differentials are allowed to exist. Ultimately, parents mostly want teachers who will be qualified and effective in educating their children. Jealousy over differential pay among teachers can be mitigated by paying superior teachers for outstanding performance regardless of field the union issue can be assuaged by the potential that higher wages for some teachers may pull up the wages for all.

However, despite the objections, these studies are evidence that there is a real need to change the current salary compensation policies in the U.S. school system. Uniform wage rates for all teachers are not an effective system and they hamper any successful effort to address the shortage of qualified teachers in specific disciplines. As one education expert noted, "It's crazy to pay the same salaries to people in high-demand subjects (e.g. high school science and math) as to those in high-supply fields (e.g. middle school social studies)" (Izumi, 2001).

In the last few years, some states in the United States including Florida and Ohio have proposed legislation to permit wage differentials for specific regional areas or disciplines. While this is a step in the right direction, without universal wage differentials, those who possess teaching skills that presently work outside of the teaching profession are unlikely to be attracted into teaching; and limited differentials may skew the compensation across school districts such that the best teachers may be concentrated in just a few schools at the disadvantage of the rest.

THE ECONOMICS OF NO WAGE DIFFERENTIALS

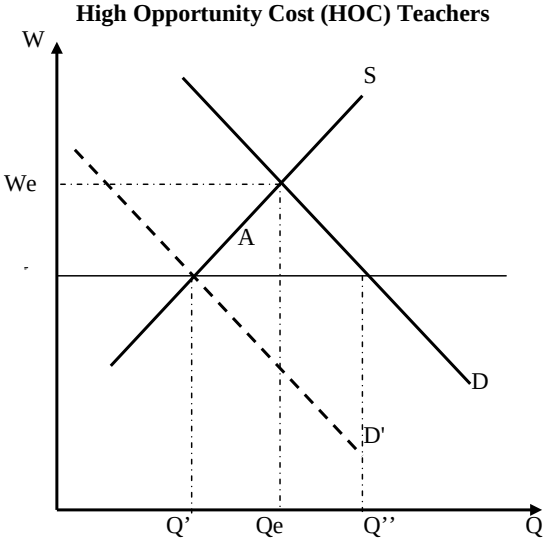
So, if wage scales are so inappropriate, why do they exist? The answer derives from two fundamental influences. The first is obvious – paying all of the teachers uniformly for the same level of experience and/or educational level is less expensive for a municipality (particularly since teacher salaries have always been very low compared to non-teacher salaries with corresponding qualifications). However, this naturally ignores the potential damage done to students who are instructed by those who perceive that they are underpaid; and thus give less than maximum effort. The second perspective is that all secondary teachers do the same thing -- teach students. However, NBA players play the same game as development league players, but their salaries differ by millions. Quarterbacks play the same game as offensive linemen, but their incomes are not remotely similar. It has always amazed the authors of this paper that those with unique skills to perform their vocations at the highest levels are virtually always compensated to do so, but those most responsible (or second most to parents) for preparing the next generation of citizens are paid woefully, and uniformly regardless of effectiveness!

In addition, universities are generally perceived to be more effective than secondary schools. Not surprisingly, universities attract the most qualified individuals to teach these advance students across disciplines, in large part due to definitive wage differentials for faculty. Like professional athletes and university professors, teachers that have different skills, and more importantly, different opportunity costs, need to be paid commensurately to maintain quality among the larger opportunity cost group.

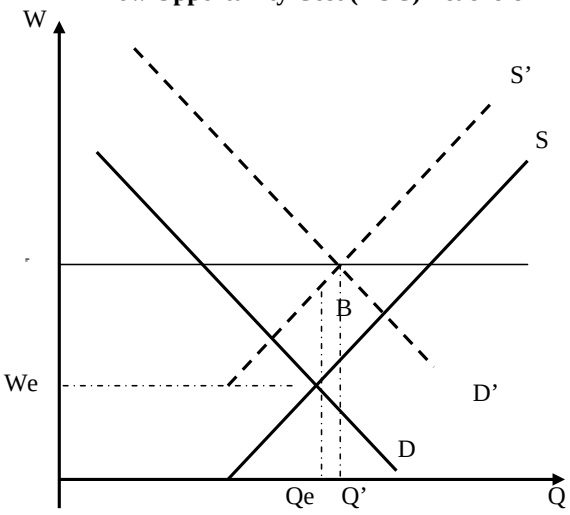
To clarify the inefficiency of the status quo, consider the graphs below. The lack of wage differentials generates temporary disequilibrium in the teaching markets (graphs 1 and 2) due to the wage controls. However, the wage limits do not influence markets for high opportunity cost (HOC) teachers (graph 1) the same way they do for low opportunity cost fields (graph 2). For those teachers in fields like math and science where there are clear shortages (graph 1), the starting wage serves as a temporary price ceiling causing an excess demand for teachers in that market (Q''-Q'). Since administrators must fill those classrooms as best they can, they may use a small number of workers from the high opportunity cost, non-teaching markets willing to accept teaching wages (not likely to be the best of the workers from graph 3). But, most likely, and more frequently, these positions will be filled by low opportunity cost teachers employed to teach out of their field. However, the achievement of equilibrium in the HOC teaching markets at the constrained wage comes most from reductions in the demand for math and science

teachers (movement to Q' and $\overline{W}$). All of these effects combined cause decreases in the supply of LOC faculty in teaching markets; raising wages in graph 2, increased supply in graph 3, and creating declining wages for these latter workers. Conversely, in the market for low opportunity cost teachers (graph 2, the Arts, English, etc.), the salary scale serves as a price floor that creates a surplus at the prevailing wage. Generating equilibrium in this market means the hiring of too many Art and English teachers (move from D to D' in Graph 2 at Q'), falling supply in graph 4, and higher wages for non-teaching, low opportunity cost workers. The absence of wage differentials generates negative welfare effects (areas A-D) due to the misallocation of resources between both teaching markets and both non-teaching markets and leads to lower quantity and quality in teaching markets that are in under supply (HOC). The dynamics between the markets exacerbate the differentials between teaching and non-teaching wages and encourage (necessitate) hiring out of field. Many teachers' unions argue that the fundamental problem is that all teacher salaries are too low in general, but the reality is that while some teacher salaries are too low, many are simply too high to avoid distorting both teaching and non-teaching markets. It is no wonder that teachers exhibit so much discontent, particularly in high opportunity cost fields, which ends up generating low quality teaching for their students.

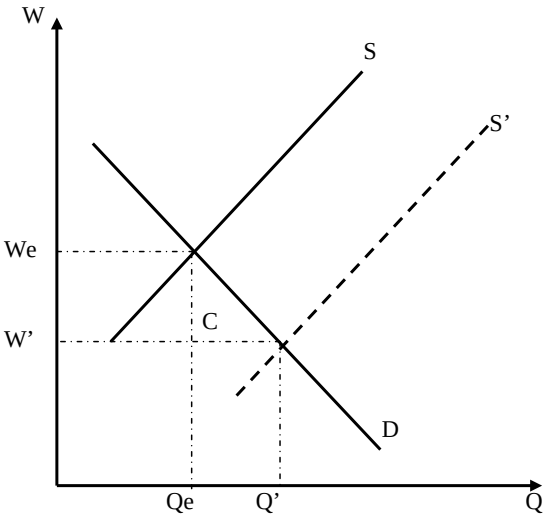
Graph 1
High Opportunity Cost (HOC) Teachers



Graph 2
Low Opportunity Cost (LOC) Teachers



Graph 3
HOC Non-Teachers



Graph 4
LOC Non-Teachers

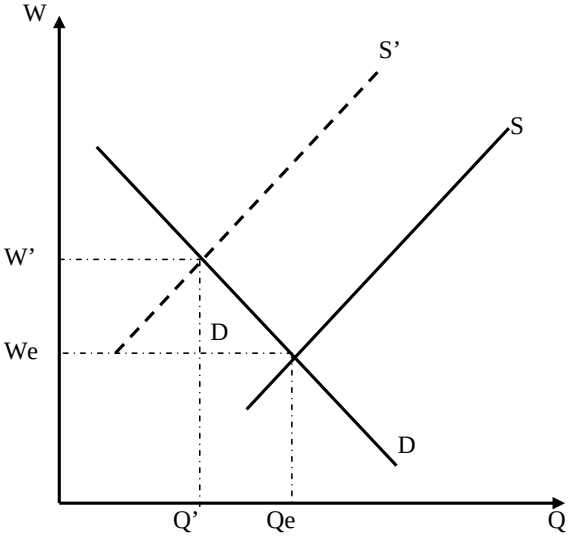


Table 1: Percentages of Public Secondary Teachers by Discipline, Various Years, 1961-2011

Fields/Years	1961	1966	1971	1976	1981	1988	1991	1991 *	1994	1994 *	2008	2008 *
The Arts	3.9	6.7	7.5	5.4	**7.9**	6.7	5.7		4.6		7.49	4.3
English	19.0	18.1	20.4	19.9	**24.8**	16.4	14.2	10.9	14.6	11.1	15.9	7.8
Foreign Languages	4.1	6.4	4.8	4.2	2.7	4.8	5.5		5.8		**5.9**	1.5
Health and Physical Education	8.2	6.9	**8.3**	7.9	6.5	5.3	5.1		4.9		6.7	2.2
Mathematics	11.4	13.9	14.4	**18.2**	15.6	13.2	12.9	8.8	14.1	9.7	13.4	4.8
Sciences	11.7	10.8	10.6	**13.1**	11.7	10.7	10.9	9.0	10.4	6.6	11.6	4.1
Social Studies	12.9	**15.3**	14.0	12.4	10.9	11.0	10.7	8.9	11.2	7.1	11.4	4.4
Special Education	0.3	0.4	1.1	3.0	2.1	8.1	8.8		**19.8**		10.2	4.7
Other	**28.5**	21.5	18.9	15.9	17.8	24.0	26.1		14.6		17.4	12.5
TOTALS	**100%**	**100%**	**100%**	**100%**	**100%**	**100 %**	**100 %**		**100 %**		**100 %**	

*Values in green were estimated from alternative sources
*Net of Out-of- Field Faculty

SOME SUPPORTING DATA

One of the fundamental realities of the graphical exposition above is that the lack of wage differentials in secondary schools should generate circumstances whereby there are too few HOC teachers in public schools and too many LOC teachers. An assessment of historical data verifies this outcome, although inferring such is complicated by educational trends and changes in relative wages. Table 1 provides data on percentages of secondary public school teachers in eight core disciplines (The Arts, English, Foreign Languages, Health and Physical Education, Mathematics, Science, Social Studies, and Special Education) individually for various years from 1961 through 2011. To make the totals add to 100%, the table also includes a category for all other disciplines designated as "Other."

Those data in the columns in regular black ink are derived from respectively, the National Education Association *Status of the American Public School Teacher* for 1961, 1966, 1971, 1976, and 1981. The remainders were developed from several of the *National Center for Education Statistics (NCES), Schools and Staffing Surveys* for 1987-88, 1990-91, 1993-94, and 2007-08. The three columns in red were derived from an April 1996 *Issue Brief* from the (NCES), Ingersoll (1999), and Feistritzer 2011.

Perusal of these data reveals a downward trend in the percentage of teachers in all of the traditional major categories due primarily to the sizable increase in the numbers of teachers who provide instruction in special education. The peak percentages by discipline arose in different years (bolded in Table 1) however all of the disciplines except special education and foreign languages occurred before 1982. Not surprisingly, as high schools have diversified their course offerings, the general trend would be expected to move away from the dominant fields of English, Math, Social Studies, and Science.

However, what is most significant for this paper results from the red columns that reflect the percentages of teachers with majors or minors in the primary field that they teach who populate public school classrooms. As is obvious, far fewer teachers with majors or minors in their teaching discipline

(or a closely related discipline, e.g., physics or computer science for math teachers, physical science for chemistry teachers) are actually teaching those subjects in our secondary schools. The unavoidable conclusion is that certified teachers from unrelated disciplines are being instilled in classrooms (e.g., physical education trained individuals teaching math, history trained teachers instructing in science). But likely in larger quantities, teachers are being recruited from outside education departments to teach a variety of courses. Since the data reported in these papers associated with the red columns include math majors teaching math, English teachers teaching language arts, etc., the implication is that non-education and non-discipline related college graduates are becoming secondary teachers in fields for which they are not qualified. In fact, programs like Teach for America and the ACE program out of Notre Dame are expressly designed to facilitate this process. According to the 2007-08 data (the most recent survey available), only one quarter of the foreign language teachers populating our classrooms would traditionally be considered qualified. Only slightly better are health and physical education teachers (2.2% relative to 6.7% for 2007-08). Now consider the 2007-08 data for the traditional dominant disciplines. While 57.4% of the teachers of Art and 49.1% of English teachers are qualified, only 38.5% of Social Studies teachers are deemed training competent, with Mathematics and Science bringing up the rear at 35.8% and 35.3% respectively. Comparing these outcomes to those for 1990-91 and 1993-94 reveals that while the trend towards fewer qualified teachers is downward for all four disciplines, the rate of decline is more sizable for math and science (i.e., the HOC fields considered in the graphs.)

The bottom line is that training of secondary teachers is less prevalent than before the 1990s. Although consideration of this dilemma dates back to what Conant called the "dirty little secret" in his 1963 landmark book, modern era scholastic institutions are putting fewer educationally qualified teachers in secondary classrooms. A naïve viewpoint on this reality cites declining state revenues allocated to education. However, ask a teacher or student from the early part of the twentieth century about their compensation for teaching prior to the 1950s; it was virtually universally below the income of other disciplines with the same level of education. The modern problem is that unlike prior to the mid-1970s, predominantly the women who were best educated had few opportunities outside of teaching to earn a level of income commensurate with their education. Today that is not the case, so to bring these women (and men) back to secondary schools will take more globally consistent compensation of teachers, commensurate with alternative employment options given their level of training and experience. Otherwise, just like a snowball rolling down a hill, the decline in educational quality will gain speed and pick up larger numbers of ill prepared young people.

HOW WILL THIS SOLVE THE PROBLEM?

Wage differentials that bring equilibrium to both categories of teaching markets independently will help bring the appropriate number of teachers by discipline to teaching within field. The need to cover math classes with PE teachers will diminish because the market will adjust to attract math teachers for these jobs. Further, the expectation of market salaries that are discipline specific will attract more students to major in math and science and presumably fewer to major in history and English, when these salaries lag behind. This is not to imply that history majors will shift to math, but that more math-oriented college students will end up teaching math instead of populating board rooms, and alternative non-academic pursuits.

However, the benefits go further. With more math oriented individuals who are drawn to teaching because of their aptitude for instruction in school classrooms, more students will likely be enlightened to the joys of learning math, leading to more math teachers in subsequent generations at all levels (including the primary level where the present hatred and ill preparation begins). Presently, our system is stuck in a downward spiral of inferior teaching begetting poor student preparation, generating even worse teachers, etc.

It is common sense that students motivated in their classrooms learn better. Attracting better teachers can only help this development process move students in the right direction, which should promote better discipline while promoting better learning. Parents show greater respect for teachers who

motivate and improve their children. Consequently, parental support for schools and the teachers will expand and perhaps even return to the levels before the 1970s. Better teachers mean fewer problems for administrators permitting more efficient allocations of scarce resources.

However, it is also important to infer that other changes need to be made for teaching to become an attractive profession for those best qualified to teach. School administrators need to adjust their perspectives regarding who the best teachers are from evaluating education coursework to evaluating field-specific backgrounds. Command of language, mathematics, and other general knowledge need exist in all teachers regardless of level or discipline if the overall ranks of instructional skills are to improve. Both men and women of superior ability and training need to replace the caretakers that populate too many classrooms if we are to improve the reputation of schools, and thus, the future viewpoint of children and parents relative to the value of education. Like in Finland, once teachers are attracted to the classroom, they will only remain if they are given the freedom to conduct their classes in the manner they deem most effective. Already, the seeds of the decline in the Finnish system exist because they too lack wage differentials by discipline. Falling relative pay issues are draining their best from the classrooms as their population ages (http://www.oecd.org/dataoecd/43/15/5328720.pdf). In the United States and elsewhere, once the right people are in these classrooms, it will be more appropriate to provide the freedom for the teachers to excel. Naturally, time will be required for better teachers to inspire students and parents to value education so that negative peer pressure, poor home environments for educational excellence, and class size issues can be overcome. However, there is no denying that under-qualified teachers (both in teaching methodology and in their discipline) who cannot motivate their charges and those students' parents, will not succeed in improving our schools.

CONCLUSIONS

This paper recommends interjecting wage differentials into the process for setting salaries for secondary teachers. High school administrators who wish to recruit effective and well qualified teachers to their classrooms will only be able to do so if those individuals can be attracted away from private industry and other government employment. Teachers are consumers like everyone else, and as employees they want to be paid commensurate with their abilities and at or beyond their opportunity costs. The present salary scale procedure used in most schools fails to generate this outcome. United States public and private schools would benefit dramatically from more efficient and more equitable payment methodologies. If Finnish schools are the best in the world because they promote fair compensation and freedom, then the United States educational bureaucracy should learn from the best and adopt a viable procedure to achieve teacher improvement -- wage differentials. Ultimately, this realization is nothing more than applying basic microeconomics.

REFERENCES

Alvarez, L. (2004). Educators Flocking to Finland, Land of Literate Children. *New York Times.* April 9, 2004.

Ballou, D. (2000). Teacher Contracts in Massachusetts. *Pioneer Institute for Public Policy Research – White Paper.* No. 12, 1-59.

Baugh, W. and Stone, J. (2002). Mobility and Wage Equilibrium in the Educator Labor Market. *Economics of Education Review* 2(3), pp. 253-274.

Baugh, W. and Stone, J. (2011). Mobility and Wage Equilibrium in the Educator Labor Market. Division of Educational Policy and Management, College of Education, University of Oregon. 48 pages.

Chingos, M. and M. West. (2012) Do More Effective Teachers Earn More Outside of the Classroom? *Education Finance and Policy.* Vol. 7, No. 1 January 12. pp. 8-43.

Conant, J. (1963). *The Education of American Teachers.* New York: McGraw-Hill.

Feistritzer, E. Profile of Teachers in the U.S. 2011. (2011). *National Center for Education Information.*

Ingersoll, R. M. (1999). The Problem of Underqualified Teachers in American Secondary Schools. *Educational Researcher.* Vol. 28, No. 2. Pp.26-37.

Issue Brief. (1996). Are High School Teachers Teaching Core Subjects Without College Majors or Minors in Those Subjects? *Office of Educational Research and Improvement.*

Levitt, S. D. and Dubner, S. J. (2005). Freakonomics: A Rogue Economist the Hidden Side of Everything. New York: HarperCollins Publishers, Inc.

ATTRACTING, DEVELOPING AND RETAINING EFFECTIVE TEACHERS

COUNTRY BACKGROUND REPORT FOR FINLAND. JUNE 2003. Retrieved, September 16, 2012 from http://www.oecd.org/dataoecd/43/15/5328720.pdf

Izumi, Lance. (2001). Differentiated Compensation in Teacher Salaries. *Pacific Research Institute*. October, No 76.

Leigh, A. & Mead, S. (2005). Lifting Teacher Performance. *Progressive Policy Institute*. April.

Monk, D.H. (1994). Subject-area preparation of Secondary Mathematics and Science Teachers and Student Achievement. Economics of Education Review. 13(2), pp. 125-45.

National Center for Policy Analysis. (1999). Teacher Accountability in Charter Schools, *Brief Analysis*. p. 285.

West, M. (2013). Do Math and Science Teachers Earn More Outside of Education? The *Brown Center Chalkboard*. April 17. Retrieved on 7/21/13 from: http://www.brookings.edu/blogs/brown-center-chalkboard/posts/2013/04/17-math-science-teachers-west

ASSESSING STUDENT UNDERSTANDING OF PRICE AND OPPORTUNITY COST THROUGH A HYBRID TEST INSTRUMENT: AN EXPLORATORY STUDY

Kristen Collett-Schmitt, University of Notre Dame
Ross Guest, Griffith University, Australia
Peter Davies, University of Birmingham, UK

ABSTRACT

This paper reports an exploratory empirical study of students' 'understanding' of price and opportunity cost based on a hybrid test instrument which consisted of multiple choice (MC) questions that also required students to explain their answers through a constructed response (CR). Data were drawn from three cohorts of students: two from an Australian university and one from a U.S. university. The information provided by the CRs attached to the MC questions tends to support extant evidence of problems with the effectiveness of standard MC questions in assessing 'understanding,' especially of price. Additional econometric analysis suggests that MC and associated CR questions may not be as effective at capturing achievement of sophisticated conceptions or 'understanding,' compared with alternative assessment tasks such as problems and essays.

Keywords: price, opportunity cost, assessment, multiple choice, constructed response
JEL codes: A20, A22, A29

Students' 'understanding' of opportunity cost and price has been widely regarded as central to economic education (e.g. Sevón & Weckström, 1989; Leiser & Halachmi, 2006; Salemi, 2005; Davies & Lundholm, 2012). However, as Davies (2011) points out, evidence from assessments of conceptions of price and opportunity cost leaves a great deal of uncertainty about what it means for a student to 'understand' price or opportunity cost.– This implies a problem for university lecturers evaluating test items on the economic concepts, and for high school teachers using similar test items to gauge students' progress and readiness to proceed to the next level. This is also a problem for researchers trying to judge the effectiveness of different forms of intervention in teaching.

In this paper, we aim to shed more light on the nature of these problems faced by economic educators and researchers, and in doing so, pilot a new approach to the design and analysis of assessment of economic 'understanding.' Our method is to examine students' answers to multiple choice questions, which are commonly used on tests to assess economic 'understanding,' alongside their accounts of why they considered a particular option to be correct. We gather evidence of students' conception of price and opportunity cost as expressed in these constructed responses. We also analyze students' multiple choice answers and the quality of their constructed responses as determinants of their overall achievement in the course.[1] . Our exploratory methodology reveals variation in students' 'understanding' of opportunity cost and price, which is missed by conventional approaches.

'UNDERSTANDING' PRICE AND OPPORTUNITY COST: CURRENT EVIDENCE

While Bloom's taxonomy suggests a distinction between 'understanding' and 'applying' a concept (Bloom, 1956), the research on students' 'understanding' of price does not (e.g. Thompson & Siegler, 2000; Meyer & Shanahan, 2002; Pang & Marton, 2003, 2005; Leiser & Halachmi, 2006). Regardless of whether this research has been undertaken in the social psychology, phenomenographic or another tradition, it has defined 'understanding' as *necessarily* including 'application' and based on experiences that are familiar to students. Using in-depth interviews with students, researchers have suggested four different conceptions of price which are evident in the way students talk or write about their experiences (Pang & Marton, 2005):

1. Price as a reflection of a good's intrinsic quality
2. Price as a reflection of cost of supply
3. Price as a reflection of consumers' willingness to demand
4. Price as a reflection of demand and supply

However, it would be premature to regard this fourfold categorization of conceptions of price as anything more than provisional. For example, Pang & Marton (2005) suggest a fifth category –price as a reflection of relative magnitudes of changes in supply and demand. Further, Davies (2011) proposes that this category be subdivided according to the context used to frame the conception (individual producer, market, or interactions between markets) and whether causation runs only from supply or demand to price or if causation from price to supply and demand is also recognized.

There has been less research on students' 'understanding' of opportunity cost. The available evidence is largely based on students' answers to multiple choice questions. Ferraro and Taylor (2005) devised a multiple choice question using the example of buying a concert ticket. By placing the question in the context of an experience which is likely to be familiar, they address the way in which the concept of opportunity cost is 'understood' in everyday experiences. However, this stance is compromised by the actual question which is, 'what is the opportunity cost?' That is, the question turns out to be asking for accuracy in a technical definition rather than seeking to expose the way that respondents understand opportunity cost in their everyday experiences. They report that only 21.6% of those attendees at an AEA conference who were asked the question opted for the correct answer. O'Donnell (2009) used the same question as Ferraro & Taylor (2005) with academics and students at his institution. He reports that the proportion of academics answering the question correctly increased when the sentence referring to opportunity cost was omitted, though his sample sizes were small. Whilst Ferraro & Taylor (2005) conclude that academic economists have a weak grasp of a fundamental concept, O'Donnell (2009) concludes that opportunity cost cannot be fundamental if such a high proportion of academics do not understand it. An alternative explanation is that this is simply a poorly-performing question, as judged by the standards usually employed in evaluating multiple choice questions. Regardless, this research implies that language such as 'better understanding' may be fit for the practice of teaching and learning theory rather than language of, 'does/does not understand.'

Greater integration of ideas has also been identified as a critical feature of 'understanding' within an academic domain (Alexander, 2000). Someone who can correctly answer multiple choice questions of the form, 'Which of these is the opportunity cost?' while

failing to use the idea of opportunity cost when asked about the movement of firms between markets, the benefits of international trade, or the impact of a tax on sales of a product, must have a trivial 'understanding' of the idea. This point has been central to the research on threshold concepts (Davies, 2012). For example, Davies & Mangan (2007) suggest that when a threshold concept of 'interaction between markets' is embedded in students' thinking, they are able to use this as an organizing idea when making sense of pricing problems. According to this proposition, it is therefore important to distinguish between students who frame their thinking about price solely in terms of individual producer decisions, students who frame their thinking about price solely in terms of behavior within an individual market, and students who frame their thinking about price in terms of interactions between markets. Likewise, it is important to distinguish between students who 'understand' opportunity cost as a way of thinking about individual decisions and students who 'understand' opportunity cost as a property of an economic system which underpins differences in value (as originally suggested by Von Wieser in 1891).

THE MC-CR DEBATE AND BEYOND

Our methodology in this paper bears heavily not just upon the aforementioned literature on student 'understanding, but also on the - literature on the relative efficacy of multiple choice (MC) and constructed response (CR) questions in assessing learning outcomes. MC questions have some clear advantages in that they are free from bias in grading and enable a wide range of content to be examined for a given test duration (Walstad, 1987; Saunders & Walstad, 1990, as cited in Becker & Johnston, 1999). The argument for CR questions has been largely framed within Bloom's taxonomy of educational objectives. For example, Biggs (1999) argues that MC questions are ill-suited to measuring what Bloom's taxonomy calls 'higher cognitive skills' that involve synthesis and evaluation. This criticism has encouraged researchers to compare students' scores on MC and CR (particularly essay) questions to identify differences in how the two types of questions assess student 'understanding.'

A number of early studies from the educational measurement literature found a high positive correlation between scores on MC and CR questions, leading some to argue that it does not matter whether MC and CR questions are used (Bennett, Rock & Wang, 1991; Bridgeman & Rock, 1993; Thissen, Wainer & Wang, 1994; Lukhele et al., 1994). Becker and Johnston (1999) also found a high correlation between the two test scores in their sample. However, they argue that this result does not mean that MC and CR questions are testing the same dimensions of learning but that an unmeasured factor, such as ability, strongly determines performance on both types of questions. More recent evidence adds support for differences in MC and CR questions.

In the context of economics education, Hickson and Reid (2011) provide exhaustive econometric evidence that CR questions add information about learning not included in MC questions. Buckles and Siegfried (2006) find that MC questions are not as reliable as CR questions in assessing synthesis and evaluation in the field of economics. They argue that the chain of reasoning required for this level of learning is hard to assess in a single or even sequence of MC questions. Kennedy and Walstad (1997) found that CR questions on the 1991 AP tests for micro and macroeconomics provided additional value over MC questions in discriminating students.

This literature on conceptual 'understanding' in economics implies other problems with MC questions. If we want to distinguish between students according to the sophistication of their conception (say) of price in a given context, we could create a MC question in which each

possible response reflects a conception which researchers have identified. This, however, poses a huge challenge for item construction. For example, consider the following question:

Question stem: Which of the following is the best explanation of a rise in price given situation *X*?

1. A change in the intrinsic quality of the good being sold
2. An increase in demand but no reference to movement along the supply curve
3. A reduction in supply with no reference to movement along the demand curve
4. A reference to both a shift and movement along demand and supply curves.

This question could never be considered to be an effective test of learning outcomes. One problem with this question is that it is difficult to avoid options that students might reason to be correct. For example, a student might select option A. as the correct answer because this outcome increases demand and ultimately price. Additionally, a student might logically deduce option D. to be correct because it subsumes other options.

Another challenge of using MC questions is that they require questions to be framed such that there is only one correct answer. Economics in the real world is rarely that straightforward. Since assessment has powerful 'wash-back' effects on what students think they have to learn and on what lecturers feel obliged to teach (Biggs, 2002), multiple choice questions inevitably encourage the belief that economic problems are unambiguous and have absolute solutions.

Our study goes beyond this MC-CR debate in the literature by uncovering students' varied conceptions of opportunity cost and price using a hybrid test instrument that asks students to provide a CR to explain their answer to a MC question. We examine the text of the CRs in detail in order to infer the nature of students' 'understanding,' and then relate performance on both CR and MC components to their final achievement in the course .We find neither the MC nor CR questions are as reliable an indicator of final achievement for high GPA students as for low GPA students. This we argue is because high GPA students tend to achieve more sophisticated conceptions, which are better captured by other assessment tasks.

THE TEST INSTRUMENT AND DATA

Our exploratory test instrument consists of four multiple choice (MC) questions, which are given in the Appendix. Questions 2 and 3 were on opportunity cost (OC) and questions 1 and 4 were on market price as the outcome of supply and demand.[2] Students were also asked to provide a brief written explanation of the reasoning behind their answer to each question. We will refer to these accounts as the students' constructed responses (CRs), and to the selection of answers A, B, C or D as the students' MC answers. The instrument is therefore a hybrid of MC and CR questions, the unique feature of which is that the CR questions are linked to the MC questions.

In order to assess the change in student 'understanding,' students were asked to complete a pre and posttest. They were advised that performance on the test instrument would not impact their final mark in the course. The pretest was administered before any substantive teaching had occurred, and the posttest was administered after the two concepts of OC and market price (with basic supply and demand analysis) had been formally studied. Students were not provided with any answers to the test after the pretest version, and the posttest was exactly the same as the pretest. Importantly for the econometric analysis reported below, the test marks were not included in students' final grade for the course. The test was taken by two cohorts of students

with the same instructor at one Australian university in 2011 and 2012 and one cohort of students at one university in the United States in 2012. After removing students who were absent from either of the tests, chose not to answer at least one question (including providing a CR), and did not have a GPA, the posttest sample size was 159 for the Australian cohorts (combined) and 188 for the U.S. cohort, while the pretest sample size was 309 for the Australian cohorts and 189 for the U.S.[3]

METHOD

Students' written explanations of their answer to the multiple choice questions were first categorised through qualitative analysis. In conducting this analysis, our aim was to avoid imposing a set of pre-ordained categories, although we are conscious that our familiarity with the research literature on conceptions of price is likely to have influenced our judgments regardless of our intentions. This process clearly identified the wide range of conceptions that students have of price and OC, with varying depths of 'understanding' of price in particular, which is consistent with previously-cited literature. We provide examples of our categorizations in the Appendix.

One question we wish to address is whether students' answers to these multiple choice questions provide reliable evidence about underlying conceptions. To answer this question, we cross-tabulated students' answers to the multiple choice questions with the categorizations of conceptions suggested by students' CR responses. The coding into categories, the score for each category, some examples of responses, and the number of student responses in each category are given in the Appendix. The number of student responses includes both the pretest and posttest results. Although the response categories are not, in principle, mutually exclusive, no response was assigned to more than one category in order to facilitate analysis.

After the cross-tabulation of MC and CR scores, we analyzed the following relationships:

1. The correlation between the MC and CR scores for price and the same for OC.
2. The correlation between the two MC scores for price and the same for OC; and the correlation between the two CR scores for price and the same for OC.
3. The significance of the MC scores and CR scores on both the pretest and posttest as determinants of overall achievement in the course, controlling for the students' grade point average (GPA) at the start of the course.

DATA ANALYSIS AND RESULTS

If the MC questions are good indicators of the level of 'understanding,' we would expect the MC scores on the two price questions to be highly correlated and the MC scores on the two OC questions to be highly correlated. The same should apply to the CR scores. Since both types of questions evaluate the same broad concept, we might expect the MC scores and the CR scores to be highly correlated for both the price and the OC questions (Lukhele et al., 1994; Bridgeman & Rock, 1993). The results of these correlations are given in Table 1. The results in Table 1 apply only to the posttest sample in order to investigate the effects of 'understanding' after the concepts of price and opportunity cost were addressed in the course. Table 1 indicates that, while the correlation coefficients are positive (and statistically significant), the strength of the correlation for the price questions is much lower than for the OC questions. In particular, the

correlation between the sum of the two MC scores and the sum of the two CR scores for the price questions is much lower than the corresponding scores for the OC questions.

TABLE 1: Correlation coefficients between scores on test items (posttest correct answers)

Test items	Correlation coefficient (phi)	Probability on χ^2 test
The two MC scores on price	0.20	0.01
The two MC scores on OC	0.54	0.00
The two CR scores on price	0.33	0.00
The two CR scores on OC	0.49	0.00
The sum of MC scores and sum of CR scores on price	0.68	0.00
The sum of MC scores and sum of CR scores on OC	1.02*	0.00

*Notes:
(i) Unlike the Pearson correlation coefficient, the phi coefficient is generally not bound from -1 to +1.
(ii) In order to ensure a suitable sample size, NA is treated as a separate category for all correlation coefficients.

Given that the MC and CR scores are positively correlated, albeit more for the price questions, we then explored whether CR scores add value to MC scores – that is, whether CR scores are capturing learning outcomes beyond those captured by MC scores. As a preliminary exercise, we run a simple (OLS) regression, for all four questions combined on the post test, of CR scores on MC scores, GPA, and dummy variables to capture gender and institutional effects:

$$CRscore_i = \alpha_0 + \alpha_1 MCscore_i + \alpha_2 GPA_i + \alpha_3 DFA_i + \alpha_4 DFUS_i + \alpha_5 DMUS_i + \varepsilon_i \qquad (1)$$

$$\begin{array}{cccccc} 0.03 & 2.80 & 0.20 & -0.07 & 0.11 & -0.13 \\ (0.08) & (15.06) & (3.21) & (-0.46) & (0.63) & (-0.78) \end{array} ,$$

where *CRscore* and *MCscore* are the student's average scores on the two CR and MC items respectively, DFA is a dummy taking the value of 1 for female Australian students and zero otherwise, DFUS is a dummy taking the value of 1 for female U.S. students and zero otherwise, and DMUS is a dummy taking the value of 1 for male U.S. students and zero otherwise (hence the omitted category is male Australian students). The numbers in parentheses are *t* statistics. The inclusion of GPA in this regression is based on the assumption that higher GPA students tend to achieve higher-order learning outcomes; hence the GPA variable is meant to capture any higher-order learning outcomes in CR scores not captured by MC scores. We find that GPA is statistically significant and positive, suggesting that there may be information about the degree of sophistication in student conceptions of price and OC that is reflected in the CR scores but not captured by the MC scores. Gender and institutional effects were not statistically significant.

More detailed analysis of the CR and MC responses sheds further light on the type of information captured by CR scores. Consider the two questions on price (Questions 1 and 4). In Question 1 for example, of the students who chose the correct MC response (B), 23% mentioned demand without mentioning supply and 25% mentioned supply without mentioning demand, even though the correct MC response implicitly refers to the importance of both supply and demand. For the other question on price (Question 4), of the students who chose the correct MC response (C), 30% had entirely fallacious reasoning. This is higher than the 25% of students

expected to choose the right MC response simply by guessing. What is more concerning is the fact that 13% of students who chose the wrong MC response for Question 4 gave a high quality CR, which was in the same category as the explanations given by the 70% of students who chose the correct MC response. These results suggest that the MC options are not properly discriminating the quality of 'understanding' of the concept of price. The OC questions (Questions 2 and 3) also reveal more reliability than the questions on price. For Question 2, 86% of students who chose the correct MC response also gave the highest quality CR. For Question 3, the corresponding figure was 75%. In both OC questions, less than 5% of students who chose the wrong MC response gave the best CR.

A comparison of pre and posttest responses also indicates a problem with MC questions as an indicator of learning for the concept of price. For Question 1, 8% more students chose the correct MC response on the posttest than on the pretest. For Question 4, the corresponding figure was 7%. This is a low improvement rate. The improvement rate is better for the CRs, especially for Question 1. As evidence, for each of Questions 1 and 4, only half the proportion of students gave an irrelevant or nonsensical CR answer on the posttest, compared with the pretest (5% compared with 10%). For Question 1, 13% more students gave the highest quality CR on the posttest test than on the pretest; for Question 4, however, the corresponding figure was only 2%. This suggests that assessing learning of the concept of price is complicated and difficult even using CR questions.

The improvement rate was much better for the OC questions. For Question 2, 19% more students chose the correct MC response on the posttest than on the pretest. For Question 3, the corresponding figure was 37%. And for both Questions 2 and 3, of the students who chose the correct MC answer, roughly the same proportion provided the best CR on the posttest as on the pretest. This means that the relatively strong improvement in performance on the OC questions was reflected in both the CRs and MC answers.

These results suggest that these MC questions do not pick up the wide range of 'understandings' of the concept of price that are held by all students. The MC questions are also not picking up the change in 'understanding' from the pretest to the posttest. Further, assessment of 'understanding' appears to be more difficult for the concept of price than for opportunity cost, regardless of the type of question used. However, since this is an exploratory study with a small number of items, we cannot discount the possibility that the results are peculiar to these particular questions.

Next we investigate the significance of the MC scores and CR scores as determinants of overall achievement in the course, for both the pretest and posttest. Here, we do not distinguish between the price and OC questions. Rather, we average the CR scores for all four questions and do the same for the MC scores. A student's final mark in the course is the indicator of overall achievement. For both the U.S. and Australian students approximately 50% of the final mark was determined from MC questions (including the mid-term and final exams). The remaining 50% of the final mark was determined by problem- solving tasks and short essays (and, in the case of the U.S. students, a 5% class participation mark). Since a student's final mark is not dominated by the one assessment type, there is no obvious bias in the ability of either CR or MC questions on the test instrument to determine the final mark. Again, note that scores on the test instrument were not included in the final mark.

We start with a simple regression model for each question type:

$$Final\_mark_i = \alpha_0 + \alpha_1 CRscore_i + \alpha_2 DFA_i + \alpha_3 DFUS_i + \alpha_4 DMUS_i + \varepsilon_i \qquad (2)$$

$$Final\_mark_i = \alpha_0 + \alpha_1 MCscore_i + \alpha_2 DFA_i + \alpha_3 DFUS_i + \alpha_4 DMUS_i + \varepsilon_i \quad (3),$$

where *Final_mark* is the student's final mark[4] for the course as a measure of overall achievement. There is a potential endogeneity problem due to measurement error in the right hand variables *CRscore* and *MCscore*, since these are measured at a time period prior to the measurement of the dependent variable *Final_mark* (Becker and Salemi 1977). However, we reject the null hypothesis of endogeneity applying a Hausman test to each equation, using the square of *CRscore* and *MCscore*, respectively, as instruments. This was also the case for the models that follow.

Table 2 presents the results for the posttest data and Table 3 includes results for the pretest data for the three cohorts combined. Both the CR and MC scores are highly significant (and positive) determinants of the final mark for both the posttest (Table 2, Models (2) and (3)) and pretest (Table 3, Models (2) and (3)) regressions. The institutional effect was significant in all models and the gender effect was significant in most models. It therefore turned out to be important to control for these effects but we do not discuss them further as they are outside the focus of this paper.

TABLE 2: Regression Results for Posttest Data
The dependent variable is the final mark for the course.

	(2)	(3)	(4)	(5)	(6)	(7)
CRscore	2.98**		2.33**		9.56**	
	(6.18)		(5.14)		(4.10)	
MCscore		7.12**		5.67**		40.63**
		(3.14)		(2.74)		(4.09)
GPA			5.42**	5.92**	8.92**	9.55**
			(7.83)	(8.44)	(6.85)	(7.81)
CRscore*GPA					-1.32**	
					(-3.16)	
MCscore*GPA						-6.38**
						(-3.59)
Female, Aust	-3.26*	-3.21*	-5.07**	-5.21**	-4.71**	-5.36**
	(-1.90)	(-1.80)	(-3.17)	(-3.16)	(-2.97)	(-3.31)
Male, U.S.	20.12**	20.90**	12.84**	12.70**	13.18**	13.09**
	(11.96)	(11.74)	(7.10)	(6.72)	(7.37)	(7.04)
Female, U.S.	19.26**	20.70**	12.20**	12.59**	12.74**	12.88**
	(10.98)	(11.40)	(6.59)	(6.59)	(6.94)	(6.85)

Notes: All equations were estimated using Ordinary Least Squares, with t-statistics given in parentheses. The sample size was 347 after adjustments. Significance at 5% and 10% is denoted by ** and * respectively.

TABLE 3: Regression Results for Pretest Data
The dependent variable is the final mark for the course.

	(2)	(3)	(4)	(5)	(6)	(7)
CRscore	3.20**		2.67**		3.14	
	(5.56)		(5.34)		(1.55)	
MCscore		7.42**		5.12**		0.55
		(2.98)		(2.36)		(0.06)
GPA			6.80**	6.92**	6.98**	6.58**
			(12.84)	(12.78)	(7.39)	(7.82)
CRscore*GPA					-0.09	
					(-0.24)	
MCscore*GPA						0.85
						(0.52)
Female, Aus.	-1.31	-1.41	-3.72**	-3.93**	-3.73**	-3.97**
	(-0.83)	(-0.86)	(-2.68)	(-2.76)	(-2.69)	(-2.78)
Male, U.S.	21.31**	22.36**	10.61**	11.57**	10.65**	11.39**
	(11.43)	(11.55)	(5.84)	(6.15)	(5.83)	(5.95)
Female, U.S.	21.57**	22.75**	10.89**	11.83**	10.92**	11.74**
	(11.25)	(11.67)	(5.87)	(6.25)	(5.86)	(6.17)

Notes: All equations were estimated using Ordinary Least Squares, with t-statistics given in parentheses. The sample size was 498 after adjustments. Significance at 5% and 10% is denoted by ** and * respectively.

Next we control for GPA in Models (4) and (5):

$$Final\_mark_i = \alpha_0 + \alpha_1 CRscore_i + \alpha_2 DFA_i + \alpha_3 DFUS_i + \alpha_4 DMUS_i + \varepsilon_i \quad (4)$$

$$Final\_mark_i = \alpha_0 + \alpha_1 MCscore_i + \alpha_2 DFA_i + \alpha_3 DFUS_i + \alpha_4 DMUS_i + \varepsilon_i \quad (5),$$

where *GPA*[5] is the student's grade point average at the start of the course[6] and *i* is the student subscript. *GPA* is highly significant, while the MC and CR scores remain significant in both the pretest and posttest regressions. This suggests that both assessment types are effective indicators of final achievement.

Perhaps the more interesting question is whether the strength of MC and CR scores as determinants of the final mark depends on the academic ability of the students. The effect could go either way *a priori*. Take the potential effect of GPA in mediating the effect of CR scores on the final mark. It may be that the performance on CR questions is a better indicator of final achievement for lower-ability students than for higher-ability students, because the former are more likely to be confused by the language of a MC question and therefore benefit more from the opportunity to explain their 'understanding' through a CR. On the other hand, we might observe the opposite effect if the higher-ability students are able to demonstrate their more sophisticated 'understanding' on CR questions rather than through a MC box-ticking exercise. Assuming GPA is a measure of academic ability, we introduce the interaction terms *MCscore*GPA* and *CRscore*GPA*:

$$Final\_mark_i = \alpha_0 + \alpha_1 CRscore_i + \alpha_2 GPA_i + \alpha_3 CRscore_i * GPA_i + \alpha_4 DFA_i + \alpha_5 DFUS_i + \alpha_6 DMUS_i + \varepsilon_i \quad (6)$$

$$Final\_mark_i = \alpha_0 + \alpha_1 MCscore_i + \alpha_2 GPA_i + \alpha_3 MCscore_i * GPA_i + \alpha_4 DFA_i + \alpha_5 DFUS_i + \alpha_6 DMUS_i + \varepsilon_i \quad (7)$$

The CR scores, MC scores and interaction terms are significant in the posttest but not pretest regressions, indicating that the effects of MC and CR scores are due to the learning process. In Model (6) of the posttest regression, the CR score by itself is significantly positive while the interaction variable, CRscore*GPA, is significantly negative. The same is true in Model (7) of the posttest regression for the MC score. This indicates that GPA is mediating the impact of the two types of scores on final marks. In other words, students with higher CR scores (on the posttest) achieved a higher final mark, but this effect was weaker for students of higher ability. The same is true for MC scores. There must therefore be some learning outcomes captured in final achievement that are measured less accurately by MC and the associated CR questions for high-GPA students than for low-GPA students.[7] Those learning outcomes are measured by the assessment tasks included in final achievement that are not MC or CR questions, such as the problem-solving tasks and short essays that make up 50% of final achievement. These types of tasks may be more effective in discerning more sophisticated conceptions, compared with MC and CR questions. On the assumption that high-GPA students achieve more sophisticated conceptions than low-GPA students, the results suggest that MC and CR questions are not as effective in discerning these more sophisticated conceptions as are the other tasks included in the final mark.

CONCLUSION

The exploratory hybrid test instrument analyzed here consisted of multiple choice (MC) questions that required students to construct short paragraph responses (CR) to justify their choices. The analysis of the CRs indicated a wide range of conceptions of price and opportunity cost, in particular varying depths of 'understanding' of price, which is consistent with the literature. The additional information provided by the CRs revealed shortcomings in the effectiveness of standard MC questions in assessing 'understanding' of economic concepts, especially price.

The regression analysis indicated that both MC scores and CR scores on the posttest were alone significant determinants of the final mark, but their effect also depends on the student's GPA – the effect of both MC and CR scores was weaker for higher-GPA students. Given that approximately 50% of the final mark was made up from other types of assessment tasks consisting of problems and short essays, the mediating effect of GPA suggests that performance on these alternative tasks varies with GPA. We posit the reason may be that higher GPA students achieve more sophisticated conceptions, which are more reliably captured by assessment tasks such as problems and short essays. In that case, the MC questions and associated CRs would not be as effective in capturing the sophistication of student conceptions.

These results, although exploratory due to the small sample of MC questions, suggest implications for assessment of fundamental concepts such as price and OC in the wider context of international progress towards assessment of learning standards in economics. The OECD's Assessment of Higher Education Learning Outcomes (AHELO)[8] Feasibility Study, which was completed at the end of 2012, has developed a test instrument consisting of 45 multiple choice (MC) questions and one constructed response (CR) question (not linked to any of the MC questions). It is therefore an instrument heavily reliant on MC questions. The evidence in this paper suggests caution in relying on such instruments. It provides an argument for weighting the balance further towards CR questions relative to MC questions, and supplementing with other

types of assessment tasks such as problems and essays. Of course, such decisions have to take account of the higher costs of the grading and moderating process for questions other than MC.

ENDNOTES

1. The test used to provide this evidence was not part of overall assessment in the course. The "course" refers to the 13 week program of study, which might otherwise be called a "unit" or "subject".

2. Question Q1 came from the Economic Literacy Quiz produced by the National Council of Economic Education, at http://www.councilforeconed.org/news-information/economic-literacy-quiz/; and Questions 2 to 4 came from the UK Economics Network Test Bank, at http://www.economicsnetwork.ac.uk.

3. For the data from the Australian university, a potential bias is acknowledged in that the smaller sample on the posttest test reflected the smaller number of students who happened to attend class on the date that the test was administered, since it was administered in a normal lecture (without prior notice). A number of students had decided not to regularly attend class since the time of the pretest, which was on day one of the course.

4. Final marks were determined on a 100-point scale for all three cohorts.

5. Students at the U.S. university were enrolled as first-semester students, and thus did not have a GPA at the start of the course. For these students, self-reported scores on the ACT or SAT standardized U,S, college admissions exams were used as a proxy for prior learning and ability. First, reported ACT or SAT scores were divided by the maximum possible score on each exam. Scores were then multiplied by 7, in order to keep them on the same 7-point scale as GPA at Griffith University. If students completed both the SAT and ACT, an average of the two scores was used instead. Econometric models do not distinguish between SAT and ACT scores.

6. Since GPA was obtained at the start of the course, it does not include the final mark that is being represented by the dependent variable in this regression.

7. These results may reflect a ceiling effect, in that higher ability students (who score higher grades) have less scope to demonstrate the effect of 'understanding' on their final grades simply because they are closer to the maximum marks that can be awarded for the course. However, there were a small proportion of scores in the 80's and 90's among all cohorts, which suggests that most of the high ability students had as much scope to improve their final scores as the low ability students.

8. AHELO stands for Assessment of Higher Education Learning Outcomes. The website is http://www.oecd.org/document/22/0,3746,en_2649_35961291_40624662_1_1_1_1,00.html

REFERENCES

Alexander, P.A. (2000). Toward a model of academic development: Schooling and the acquisition of knowledge: The sequel. *Educational Researcher*, **29:** 28–33.

Becker, W.E. & M. Salemi. (1977). The learning and cost effectiveness of AVT supplemented instruction: specification of learning models. *Journal of Economic Education, 8,* 77-92.

Becker, W. & C. Johnston. (1999). The relationship between multiple choice and essay response questions in assessing economics understanding. *The Economic Record, 75*(231), 348-357.

Bennett, R.E., D. A. Rock & M. Wang. (1991). 'Equivalence of free-response and multiple-choice items.' *Journal of Educational Measurement, 28*(1), 77–92.

Biggs, J. (1999). *Teaching for quality learning at university.* Buckingham: The Society for Research into Higher Education and Open University Press.

Biggs J.B. (2002). *Aligning teaching and assessment to curriculum objectives*. LTSN Imaginative Curriculum Guide IC022.

Bloom, B.S. (Ed.), M.D. Engelhart, E J. Furst, W.H. Hill & D.R. Krathwohl. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook 1: Cognitive domain.* New York: David McKay.

Bridgeman, B. & D. A. Rock. (1993). Relationships among multiple choice and open ended analytical questions. *Journal of Educational Measurement, 30*(4), 313-29.

Buckles, S. & J.J. Siegfried. (2006). Using multiple-choice questions to evaluate in-depth learning of economics. *Journal of Economic Education, 37,* 48–57.

Davies, P. (2011). Students' conceptions of price, value and opportunity cost: some implications for future research. *Citizenship, Social and Economic Education, 10*(2-3), 101-110.

Davies, P. (2012). Threshold concepts in economics. In G. Hoyt and K.M. McGoldrick (Eds.), *The International Handbook on Teaching and Learning Economics (pp. 250-256).* Aldershot: Edward Elgar Press.

Davies, P. & C. Lundholm. (2012). Young people's understanding of socio-economic phenomena: conceptions about the free provision of goods and services. *Journal of Economic Psychology, 33,* 79-89.

Davies, P. & J. Mangan. (2007). Threshold concepts and the integration of understanding in economics. *Studies in Higher Education, 32,* 711-726.

Ferraro, P.J. & L. O. Taylor. (2005). Do economists recognize an opportunity cost when they see one? A dismal performance from the dismal science. *Contributions to Economic Analysis and Policy, 4,* 1-12.

Harris. R.B. & W.C. Kerby. (1997). Performance sssessment as a complement to multiple-choice testing: the evidence from California's golden state examinations in economics. *Journal of Economic Education, 28*(2), 122-34.

Heath, J.A. (1989). An econometric model of the role of gender in economic education. *American Economic Review, 79*(2), 226-230.

Hickson, S. & B. Reed. (2011). More evidence on the use of constructed-response questions in principles of economics classes. *International Review of Economics Education, 10*(2), 28-48.

Kennedy, P. & W. Walstad. (1997). Combining multiple-choice and constructed-response test scores: an economist's view. *Applied Measurement in Education, 10* (4), 359-375.

Leiser, D. & R.B. Halachmi. (2006). Children's understanding of market forces. *Journal of Economic Psychology, 27,* 6-19.

Lukhele, R., D. Thissen & H. Wainer. (1994). On the relative value of multiple-choice, constructed-response, and examinee-selected items on two achievement tests. *Journal of Educational Measurement, 31,* 234–250.

Marton, F. & M. F. Pang. 2008. The idea of phenomenography and the pedagogy of conceptual change. In S. Vosniadou (Ed.), *International Handbook of Conceptual Change (*pp. 533-559). New York: Routledge.

Meyer, J.H.F. & M. Shanahan. (2002). On variation in conceptions of 'price' in economics. *Higher Education, 43,* 203-225.

Moss, P.A. (1992). Shifting conceptions of validity in educational measurement: implications for performance assessment, *Review of Educational Research, 62,* 229-258.

O'Donnell, R.(2009). The concept of opportunity cost: is it simple, fundamental or necessary? *Australasian Journal of Economics, 6* (1). Retrieved July 31, 2013 from http://www.uq.edu.au/economics/AJEE/docs/index.php?dir=Volume+6%2C+Number+1%2C+2009%2F.

Pang, M. & F. Marton. (2003). Beyond "lesson study": comparing two ways of facilitating the grasp of some economic concepts. *Instructional Science, 31,* 175-194.

Pang, M. & F. Marton. (2005). Learning theory as teaching resource: enhancing students' understanding of economic concepts. *Instructional Science, 33,* 159-191.

Pong, W-Y. (1998). Students' ideas of price and trade. *Economic Awareness, 9*(2), 6-10.

Salemi, M. (2005). Teaching economic literacy: why, what how? *International Review of Economics Education, 4*, 46-57.

Saunders. P. & W. Walstad. (1990). *The principles of economics course: handbook for instructors.* New York: McGraw Hill.

Sevón, G. & S. Weckström. (1989). The development of reasoning about economic events: A study of Finnish children. *Journal of Economic Psychology, 10,* 495-514.

Shanahan, M., G. Foster & J.H.F. Meyer. (2006). Operationalising a threshold concept in economics: a pilot study using multiple choice questions on opportunity cost. *International Review of Economics Education, 5,* 29-57.

Tan, K.C.D., N.K. Goh, L.S. Chia & D.F. Treagust. (2002). Development and application of a two-tier multiple choice diagnostic instrument to assess high school students' understanding of inorganic chemistry qualitative analysis. *Journal of Research in Science Teaching, 39,* 283-301.

Thissen, D., H. Wainer & X. Wang. (1994). Are tests comprising both multiple-choice and free-response items necessarily less unidimensional than multiple-choice tests? An analysis of two tests, *Journal of Educational Measurement, 31,* 113–23.

Thompson, D. R., & R.S. Siegler. (2000). Buy low, sell high: The development of an informal theory of economics. *Child Development, 71,* 660–677.

Von Wieser, F. (1891). The Austrian school and the theory of value. *Economic Journal, 1,* 108-121.

Voska, K.W. & H.W. Heikkenen, (2000). Identification and analysis of student conceptions used to solve chemical equilibrium problems, *Journal of Research in Science Teaching, 37,* 160-176.

Walstad, W.B. (1987). Measurement instruments in William E. Becker and William B. Walstad (eds). *Econometric Modeling in Economic Education Research,* Boston: Kluwer-Nisjhoff.

APPENDIX

THE HYBRID TEST INSTRUMENT

Answer all questions and underneath each question briefly explain your answer.

Question 1:

If your local government sets a maximum amount that landlords can charge in rent, what is the most likely result?

A. There will be more apartments available than people want to rent.
B. There will be fewer apartments available than people want to rent.
C. The number of apartments available will be equal to the number of people that want to rent.
D. Don't Know.

Please explain:

_____
_____
_____
_____
_____
_____
_____

Question 2:
A ticket in a raffle in which the prize is a day trip to Sydney is purchased for $1. The regular price for this trip is $200, but the organisers of the raffle are offering a cash sum of $150 to the winner as an alternative to the trip itself. What will be the real cost for the winner of the raffle, if they decide to go on the trip?
A. $1
B. $150
C. $199
D. $200

Please explain:

_____
_____
_____
_____
_____
_____

Question 3:
If people are offering to pay $200 for tickets to the World Cup final on the black market, and someone gives you a ticket, what does it cost you to attend the game ?
A. $200 because you could have sold the ticket for $200 on the black market.
B. Nothing because you have not paid anything for the ticket.
C. $200 less the value of your time if you did not go to the game.
D. $200 plus the value of your time if you did not go to the game.

Please explain:

_____
_____
_____
_____
_____
_____
_____

Question 4:
Suppose a massive new oil field was discovered, causing the price of petrol to fall to $0.40 per litre for the foreseeable future. According to economic analysis, which of the following would most likely happen?

A. The price of small, fuel-efficient vehicles would probably rise as suppliers were forced to cope with falling demand.
B. The price of electric cars would fall, but only if they could be designed to run on petrol instead of electricity.
C. Prices for vehicles with relatively low fuel efficiency, such as SUV's (or 4 wheel drives), would probably rise.
D. The price of electricity would probably rise.

Please explain:

_____
_____
_____
_____
_____
_____
_____

END OF TEST

QUALITATIVE DIFFERENCES BETWEEN STUDENTS' CONSTRUCTED EXPLANATIONS OF THEIR ANSWERS

These data encompass responses from all three student cohorts on both the pretest and posttests.

QUESTION 1			Multiple choice option and numbers of students. * denotes correct answer				
Type	**Score**	**Example**	No response or multiple choice answers	A	**B***	C	D
No response	NA		10	6	14	5	10
0. Irrelevant or nonsensical answer	0		0	56	46	20	12
1. Explanation of effect on price that is unrelated to supply and demand	0	"The government rent limitation only sets the maximum price a landlord can't set for rent, and does not necessarily effect the 'wanting to rent' effect of people."	1	30	30	20	9
2. Maximum price will have no effect	0	"Whether or not the government impose maximum amount of rent restrictions or not will not have a direct impact on any of the above answers."	0	2	4	12	1
3. Maximum rent will increase demand	1	"It will increase the popularity of renting. This is because people will become more financially able to rent making the practice more popular. More demand, same supply."	0	7	187	9	2
4. Maximum rent will reduce supply	2	"The landlords who are trying need to charge more than the max amount set by the government will no longer want to/be able to support/invest in the apartment. New potential landlords will be less inclined to enter the market due to lack of freedom in terms of rent they charge."	0	8	125	1	2
5. Maximum rent will increase demand and reduce supply	4	"If the government cap the amount that can be charged for rent I think this will make less apartment available. If the amount is capped to a reasonable price people would be able to afford a more range of apartments meaning more people will be looking for apartments."	0	2	77	0	0
6. No further consideration of alternatives	2	"The amount of information is insufficient to arrive at an answer. Ideally we would also like to know whether the amount the governments sets is higher or lower than the current occupancy rate for available houses and the current number and more info of new dwellings being built."	1	2	3	3	25

QUESTION 1			Multiple choice option and numbers of students. * denotes correct answer				
Type	**Score**	**Example**	No response or multiple choice answers	A	**B***	C	D
7. If maximum rent is set above the current equilibrium then there will be a glut (S>D) (i.e. misunderstands equilibrium process), but if max is price below equilibrium there will be a shortage (D>S)	4	"It depends on whether this would encourage landlords to raise or lower the price in comparison to the previous market value. However, it would be most likely that the government would impose this law if the rents were too high. Therefore, the rents would be lower and there would be fewer apartments available."	0	5	11	2	14
8. If maximum rent set below equilibrium there will be a shortage, but otherwise no effect	5	"If there is a maximum set (assuming that the maximum is lower than the landlords preferred amount) then renting will become cheaper and therefore more people are more likely to rent instead of buying therefore creating fewer apartments."	2	4	60	1	4

QUESTION 2			Multiple choice option and numbers of students. * denotes correct answer				
Type	**Score**	**Example**	No response or multiple choice answers	A	**B***	C	D
No response	NA		17	24	19	16	6
0. Irrelevant or nonsensical answer	0		0	7	8	5	3
1. Uses the value of the item as the guide to cost	0	"Winner $1 cost to win - $200 - $1 = $199"	0	14	2	60	25
2. Treats 'real cost' as expenses directly incurred	1	"$1 as they are receiving a $200 trip for purchasing a $1 ticket. $1 is all they have spent"	0	321	3	6	0
3. Understands opportunity cost in terms of foregone earnings, but ignores the cash	3	"They put in $1 to start but it also depends if they go on a day they would normally work. Then they would lose a day's wage or a day's leave."	4	9	9	1	3
4. Understands opportunity cost in terms of foregone cash	5	"The alternative to the trip was $150 cash and so the winner is effectively spending this amount by going on the trip and not cashing in the money"	0	31	249	0	3

QUESTION 3			Multiple choice option and numbers of students. * denotes correct answer				
Type	**Score**	**Examples**	No response or multiple answers	A	B	C	**D***
No response	NA		11	14	45	22	27
0. Irrelevant or nonsensical answer	0		2	5	10	7	10
1. Treats 'real cost' as the price that is paid (says answer is 0)	1	"Because you paid nothing for the ticket therefore no cost incurred."	1	4	157	2	4
2. Recognizes that foregone wages are part of opportunity cost	3	"Opportunity cost means calculating implicit costs as well. For example you may have had the opportunity to spend your time earning money at a job but didn't because you went to the game."	1	5	16	28	39
3. Recognizes that a trade-on value of an item is part of the opportunity cost	3	"Opportunity cost = $200 could have got from selling ticket."	1	77	18	11	16
4. Recognizes both foregone earnings and trade on value in calculating opportunity cost-	5	"I believe it should be $200 PLUS the value of your time if you didn't attend the game. You could have sold the ticket for $200 which means you are $200 worse off PLUS the value of your time. You could have been working and earning money or spending time doing something you enjoy more."	3	5	3	13	288

QUESTION 4			Multiple choice option and numbers of students. * denotes correct answer				
Type	**Score**	**Examples**	No response or multiple answers	A	B	**C***	D
No response	NA		34	44	26	52	5
0. Irrelevant or nonsensical answer	0		4	24	24	36	19
1. Suggests that demand for car types depends on innate qualities of car	1	"Because people would prefer to go for cheaper option or just a petrol car."	1	11	8	7	6
2. Believes that a fall in demand with downward sloping S that does not shift will raise price	2	"The demand for fuel efficient cars may fall as three is less stress from petrol prices therefore the supply would also need to be less but at a higher price to achieve profits in which were achieved when there was a higher demand."	1	95	5	13	4
3. Expects rise in demand and price for all types of car	3	"People will probably demand more vehicles, meaning price will also increase."	0	14	2	13	0
4. Expects rise in demand for SUVs causing a rise a price (recognizes relative price)	4	"When the price of petrol falls the fuel cost for a car owner will decrease. More people would like to buy the low fuel efficiency cars since they will have less cost than before. When the demand is larger the price will be higher."	2	36	25	330	4

THE SYLLABUS EVOLVED: EXTENDED GRAPHIC SYLLABI FOR ECONOMICS COURSES

Katherine M. Sauer, University of Colorado
Lauren Calimeris, St. John Fisher College

ABSTRACT

An extended graphic syllabus is an artifact that contains limited jargon within a visual representation of a traditional course outline (i.e. a classic graphic syllabus) and incorporates visual elements into other parts of the traditional syllabus (e.g. course policies, assessments) in order to hold and focus the students' attention on the document. Its purpose is to give students a sense of the course's "big picture", to make their initial encounter with the course topics more meaningful, and to induce them to actually read and reference the syllabus throughout the semester. In addition to providing rationale for using extended graphic syllabi, this paper provides examples of classic graphic syllabi, describes how to create them, and offers examples of elements of extended graphic syllabi. Student perceptions of traditional and graphic syllabi are assessed.

Keywords: syllabus, economic education

JEL Classification: A22

A course syllabus is a staple in undergraduate education. Its fundamental purpose is to communicate the instructor's course design to students (Carnegie, n.d.). Practically speaking, it details the course topics, learning objectives, assessment and grading practices, schedule and due dates, as well as general course and school policies. The syllabus serves as a student's first introduction to a course and/or instructor and in many cases also serves as the student's first encounter with the discipline. Students' perceptions of the syllabus inform their first impressions of the course, which may linger throughout the remainder of a semester (Harnish & Bridges, 2011).

Syllabi tend to be laden with discipline-specific jargon that is either incomprehensible or meaningless to students at the start of the course (Nilson 2007). This is especially true in economics, which has both discipline-specific jargon and a prevalence of terms that are commonly used in everyday language but have a different, specific economic meaning (e.g. utility). Nettleship (1992) compared learning economics to a jigsaw puzzle. He argued that it is difficult to see the whole picture, but one needs to see it to understand where one is going. Each concept in economics is like a puzzle piece; students study the "color, shape, and size of each jigsaw piece. It is only when the students start to put the pieces together and to see the interrelationship they begin to appreciate the nature of the subject." (Nettleship, 1992, p. 69). After reading a syllabus, the introductory student is unlikely to see the "big picture" at that point. Furthermore, the student's first impression of the discipline and course may very well be confusion or overwhelm, and that negativity may persist throughout the course.

While a student's perception of a syllabus may set the tone for her semester, this does not mean the student has actually read the syllabus. To the chagrin of instructors everywhere, many students appear to have never read the document. Students are reading less in their daily lives (Nilson, 2007), and they entertain themselves with television, movies, social media, texting, video messages, and video games. Reading a multi-page, text-filled syllabus may well be a daunting task for some students. Also frustrating for instructors is using valuable class time on the first day to carefully go over the syllabus, only to have students display "syllabus amnesia" when it comes to course policies and due dates.

To give students a sense of the "big picture", to make their initial encounter with course topics more meaningful, and to get them to actually read and reference the syllabus throughout the semester, we advocate the use of an extended graphic syllabus. A classic graphic syllabus is a document that incorporates visual elements into the communication of material found in a traditional course outline. Nilson (2007, p. 26) defines a graphic syllabus as "a flowchart or diagram that displays the sequencing and organization of major course topics through the semester". It is a type of concept map, mind map, or advance organizer for the major topics of the course. Concept maps typically visually represent the hierarchical relationships within a discipline or subject (Willerman and Macharg 1991). Joseph Novak, the developer of the concept map, described them as "graphical tools for organizing and representing knowledge" (Novak & Canas, 2008, p. 1). They include concepts, which are typically enclosed in boxes or circles, with lines connecting the concepts delineating the relationships between them. The lines often have text written on them to describe the relationships (Novak & Canas, 2008). In a graphic syllabus, special attention is placed on illustrating how the course content fits together, which typically is not apparent to a discipline novice.

In this paper, we broaden the classic definition of a graphic syllabus. The *extended graphic syllabus* is an artifact that contains limited jargon within a visual representation of a traditional course outline (i.e. a classic graphic syllabus) and incorporates visual elements into other parts of the traditional syllabus (e.g. course policies, assessments) in order to hold and focus the students' attention on the document. Extended graphic syllabi may benefit both the student and the instructor. Some of the benefits include: reaching nonverbal learning styles, teaching students learning tools, encoding knowledge for long-term memory, revealing topical interrelationships within the "big picture", tightening course organization, and releasing faculty creativity (Bikitimirov & Nilson, 2003). An additional benefit for the instructor is that this type of syllabus may improve student perceptions of the course and therefore faculty evaluations. These benefits will be detailed in the Rationale section.

The remainder of this paper is organized as follows: we provide an example of classic graphic syllabi for economics courses, offer components of extended graphic syllabi, describe how to create them, present rationale and benefits for using extended graphic syllabi, and discuss the authors' experience with using extended graphic syllabi in their classes and student perceptions of the syllabi.

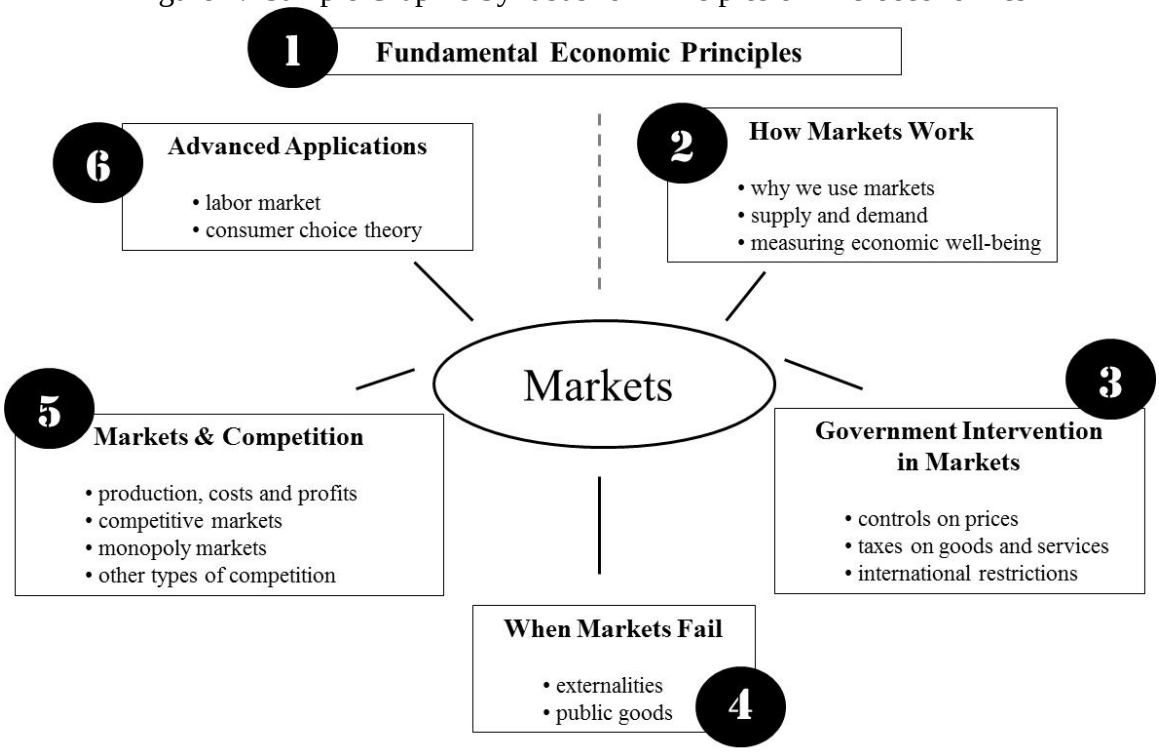
EXAMPLE OF CLASSIC GRAPHIC SYLLABUS FOR ECONOMICS COURSES

Classic graphic syllabi incorporate visual elements into the communication of material found in a traditional course outline. They avoid the use of jargon and emphasize how the course topics fit together. Each graphic syllabus will be different. The graphic syllabus should illustrate

how the instructor, the discipline expert, makes links between the concepts. This will guide the discipline novice to see the connections and the "big picture".

Figure 1 illustrates an example of a graphic syllabus for principles of microeconomics (additional examples can be seen at http://tinyurl.com/GraphicSyllabusExamples). At a glance, students can see the major course topics, how they fit together, and the approximate course schedule. The word "markets" is clearly displayed in the center, indicating the emphasis of the concept in the course. Nodes detailing various aspects of markets surround it: how they work, when they fail, government intervention, and competition. Large numbers indicate the sequence throughout the course. Relatively little jargon is used. Notice that the term "elasticity" is not present in the diagram. Elasticity is certainly covered in this class, but the term has no meaning to the discipline novice. The diagram indicates two chapters are devoted to supply and demand; omitted is the fact that the second chapter is elasticity. Not all jargon is avoidable, as evidenced by "externalities" appearing in Figure 1's Box 4; however, the larger caption of Box 4 is "When Markets Fail", which is language the discipline novice will understand.

Figure 1. Sample Graphic Syllabus for Principles of Microeconomics

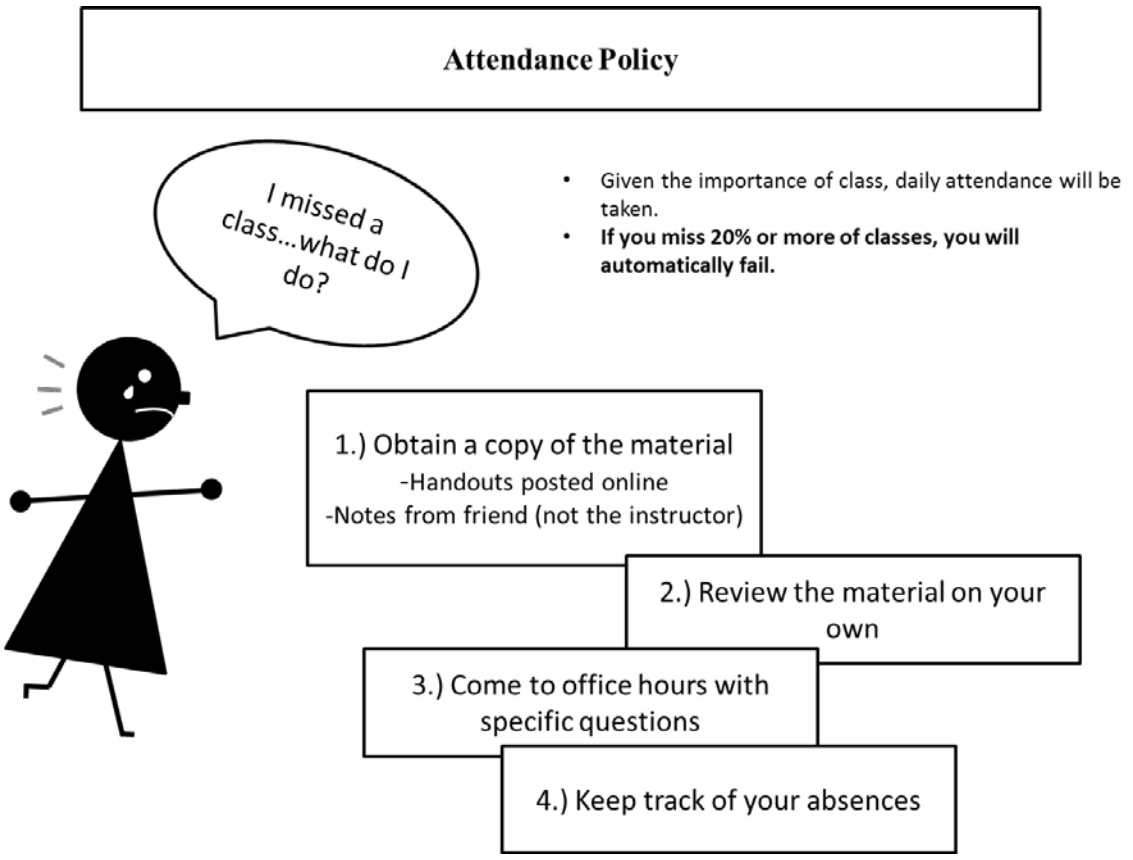


THE EXTENDED GRAPHIC SYLLABUS

There are many ways to extend the graphical components to other portions of the syllabus. The goal of an extended graphic syllabus is not to replace every single instance of syllabus text with a graphic. It is to incorporate graphic elements into the syllabus in order to focus the students' attention as they read it, to help them remember the different components, and to make it easier to reference throughout the semester.

Course-specific policies, such as an attendance policy, may have key roles in the culture of the classroom, but they may end up being glossed over in the reading of a long syllabus. Figure 2 illustrates an example of drawing attention to an attendance policy. While the policy would normally be about a paragraph of text, note that this graphic clearly outlines the step-by-step procedure students should follow if they miss a class. It clearly indicates the consequences of missing too many classes, but the graphic as a whole is meant to convey understanding and sympathy on behalf of the instructor.

Figure 2. Sample Extended Graphic Syllabus for Course Policy



The grade weights for different course components are a natural choice for a more graphic treatment. While a traditional syllabus may list the weights of the different components, seeing the grade in a graphical manner clearly shows the relative weights of the different components, especially for non-verbal learners or those unfamiliar with weighting mechanisms. Consider stacked bar charts, pie charts, or repetition of different shapes.

Graphic elements can be added to highlight any aspect of the course that the instructor wants to draw attention to. Learning objectives and course prerequisites are good candidates for this type of treatment. Learning objectives can be illustrated in flow-chart manners, showing how the objectives influence one another, or they could be incorporated into course outline or description from the classic graphic syllabus.

CREATING A GRAPHIC SYLLABUS

The process for creating a classic graphic syllabus is straightforward and can be accomplished in a few hours' time. Begin by writing down the main topics or themes that the course includes, using as little jargon as possible for introductory courses. It might be helpful to pretend you are describing the essence of your course to a non-economist friend. Focus on the big picture, not the details.

Look back at the course outline contained in your existing syllabus. Map the main topics/themes you have identified back to the detailed course outline. As an example, Table 1 contains the list of topics from one author's original principles of microeconomics syllabus as well as how she categorized the major topics in plain language. There is no "right" way to view the course topics. The purpose is to communicate your understanding of the course topics to your students. To one author, the principles of microeconomics course is about markets: students learn how markets work, mechanisms for and effects of government intervention in markets, market failures, and firm behavior in markets. To another author, the emphasis is on households and firms interacting in and creating markets. Neither view is "right," and neither view is "wrong". These views represent the individual instructors' approaches to the course. What is important is that you give your students a "road map" for the course that you teach.

Table 1
Organizing List of Topics from Traditional Syllabus for Graphic Syllabus

Original List of Topics from Syllabus	Author's Classification of Topics	Final Title in Syllabus (Figure 2)
Intro to Economics/Models		
Specialization & Trade Supply & Demand Elasticity Efficiency of Markets	Markets: Why we use them & how they work	How Markets Work
Price Controls Taxes Quantity Controls International Restrictions	Government Intervention in Markets	Government Intervention in Markets
Externalities Public Goods	Market Failure	When Markets Fail
Production Perfect Competition Monopoly Monopolistic Competition Oligopoly Factor Markets	Firm Behavior in Markets	Markets & Competition
Consumer Choice		Advanced Applications

At this point, you might identify topics or concepts that feel "tacked on" or you might identify areas that you would like to expand upon. In one microeconomics course, "consumer choice" is an example of a "tacked on" topic (Table 1). If the topic is department-mandated, take some time to brainstorm how to more fully integrate it into the course. If the topic seems out of place to a discipline expert, a discipline novice cannot be expected to make the connection between the topic and the course as a whole. In the case of a required consumer choice topic in a

principles course, one author chose to include an "Advanced Applications" category the graphic syllabus (Figure 1). When discretionary topics feel disjoint, consider elimination to free up time for more depth on other topics. This process will allow you to tighten up and better organize your course.

On scratch paper, sketch out some diagrams of how the main topics fit together. There are many ways to illustrate the various connections between course topics. Experiment! Try a variety of forms and see which are most appealing to you. Play with font (style, size, bold, italics), object shapes (rectangles, circles, stars, other polygons), line styles (arrows, dashed, thickness, curved or sharp corners), alignment (centered, linear, symmetric) and even colors (although this might be lost when you photocopy your syllabi, it would still be visible when posted as a digital file in the course learning management system). With a basic diagram in place, decide whether or not to add other information (e.g. simple descriptions, chapters, time lines, learning objectives).

Once the course description or course outline has been diagrammed, adding graphics to other parts of the syllabus is a natural extension: grading, learning objectives, other policies. To create graphics for grading, consider bar charts, pie charts, or other diagrams. Think about if you want the descriptions of the assessments along with the weighting, or if you prefer just the grade components and weights. Decide whether or not to add in other information (e.g. due dates, weights, policies). Question whether learning objectives feed into one another or if they part of a collusive whole. Determine if there are other policies that you believe would benefit from a more visual representation as opposed to a paragraph of text. Consider other policies that may be clarified with a graphic. Making the graphic "fun" with pictures may help show your sense of humor and/or assuage severity for sensitive policies (Figure 2).

Create an electronic version of your graphic syllabus either by scanning your hand-drawn diagram or by using the features of a word processing software. We created our graphic syllabi using standard PCs running Microsoft Word and PowerPoint. It can be helpful to create your diagram in PowerPoint, highlight each item in the diagram (Ctrl + A), copy the diagram (Ctrl + C), and then use "Paste Special" and "Picture (Windows Metafile)" to paste the image as a whole into your Word document. You can then drag the edges of the image of the diagram to resize it.

RATIONALE

Extended graphic syllabi may benefit both the student and the instructor. The graphical nature enables instructors to reach different types of learners and helps the students to retain information by enabling learners to visually see the interrelationships between topics. Instructors, in thinking more critically about their courses, are able to tighten the course organization and be more creative. Instructors may also benefit from improved student perceptions which translate onto faculty evaluations.

Individuals have different preferences for how they receive information and different styles of processing it. One way of categorizing the modes of taking in information is with Fleming's VARK questionnaire which measures preferences for visual, aural, read/write, and kinesthetic sources of information (available free online at www.vark-learn.com). In the VARK, the visual modality refers not to pictures but to "graphics" like relational diagrams. Students who have taken the VARK questionnaire (n = 40,674) chose an option corresponding to the visual preference 20.5% of the time and read/write 26.4% of the time (Vark-Learn, 2011). This suggests presenting the course information in a combination of text and graphics would reach a

significant number of our students in a way they prefer. With an increasing number of online courses as well as numbers of nontraditional students and nonnative speakers enrolled in courses, having a syllabus that reaches a variety of learners is increasingly important.

At many institutions, the principles courses are "service" courses that are taken by a wide variety of students, many of whom are not business nor economics majors. Allgood et al. (2004) surveyed non-recent graduates about what they remember studying in college economics courses. Students with non-business majors report remembering far fewer topics than the economics and business majors did. It is important to help all students learn topics for the long-run but perhaps especially so for the students who will take only a single economics course in their lifetime.

The use of limited-jargon graphic syllabi may help improve the long-term retention of material from our introductory courses for such students. Learning theory suggests that all people learn and process verbal and visual information in different cognitive systems (Nilson, 2007; Vekiri 2002). Graphics allow learning through dual-coding, visual learning, and conjoint retention. The essential idea behind these mechanisms is that people are able to interpret pictures and graphics and retain and comprehend more complex information from these visual aids than by reading. There are fewer "cognitive transactions" when processing a picture than when reading, because the mind processes the picture as a whole unit, while it processes individual words independently and consecutively (Vekiri, 2002). Clark and Paivio (1991) provide a review of the research on dual coding theory. Wills and Ellis (2010) review the extensive research on graphic organizers and conclude they help with dual encoding of information.

Learning is a complex process, and students do not learn new material in isolation from other knowledge. Instead, learners integrate new material with knowledge they already have (Ausubel 1960, 1963; Nilson 2007, 2010). Essentially, to learn new information, the mind needs a "big picture" in which it can put material into a context (Nilson, 2010). The graphic syllabus delivers the "big picture" to students, providing them with the context in which the new material can be stored, as concept maps provide students with the ability to assimilate new knowledge into their existing knowledge (Baitz, 2009). By seeing the "big picture" of the course on the first day of class, students are better able to visualize and to understand what they are learning and where they are going in the course as well.

As discipline experts, instructors readily see how individual course topics relate to one another. As discipline novices, our students need guidance to see the relationships. Referencing the graphic syllabus when a new topic is introduced can provide concrete context for how the topic relates to what has been learned and what is still to come; it can serve as both a review and preview.

Instructors may use the graphic syllabus as the students' first introduction to the concept mapping or mind mapping study tool. Students will then have this example at the very beginning of the class for reference. Instructors can implement concept maps either as an active learning activity done in groups during class, or they could teach it as a study tool for students to use independently for studying for exams.

Research indicates that graphic organizers improve student learning and retention of new and unfamiliar material (Ausubel, 1960; Horton et al., 1993; Manoli & Papadopoulou, 2012; Willerman & Macharg 1991). The use of teacher-prepared concept maps in the classroom significantly improves student learning (Horton et al. 1993; Willerman and Macharg 1991). Economics students responded positively to the use of concept maps in Principles of

Microeconomics courses (Marangos 2003; Marangos & Alley 2007). Budd (2004) describes specific techniques for using mind maps in economics courses.

Instructors may benefit from creating a graphic syllabus as it helps them to re-focus on the course material and to think about it in a different manner. When writing a graphic syllabus, an instructor must explicitly think about the interrelationships of concepts taught in a manner in which she may not have thought about them before. The main goal is to convey the relationships clearly in a visual manner to students. By creating a graphic syllabus, the instructor is allowing the students to glimpse how an expert relates and visualizes the concepts in the course. To create the graphics, the instructor must take a step back and analyze what she is teaching, why she is teaching it, and how it relates to the rest of the material and the course in general. Through this process, instructors may realize they would like to reorganize the course, eliminate some material, and add in other material to address particular learning objectives. In such a way, an instructor may "tighten" her course. This process also helps to keep the instructor interested in the material and to breathe new life into the course, especially if it is a course the instructor has taught multiple times. Both of your authors found the act of creating a graphic syllabus to be very helpful in this regard.

Comments your authors have received from participants in a graphic syllabus workshop indicate that creating a graphic syllabus is a pleasurable experience. It is a way to "breathe new life" into the task of revising a syllabus for the coming semester. Both of your authors enjoyed creating their extended graphic syllabi and actually look forward to making revisions to syllabi each semester. It is a chance to be expressive in an otherwise dull document.

Finally, it is important for the syllabus to set the desired tone for the course, and the syllabus influences students' perceptions about a course. Students perceived instructors whose syllabus was written in a warmer tone to be warmer, to care more about their students' learning, to be more approachable, and to care more about teaching. These impressions may influence a students' course evaluation at the end of the semester (Harnish & Bridges, 2011). Instructors who want to be perceived as caring about student learning and teaching, and who desire better course evaluations, may find the investment of transforming their syllabus will help them in this capacity. Since the graphic syllabus is a relatively new concept, early adopter advantages are likely to be present and persist for some time.

DISCUSSION

In our courses, we present the graphic syllabus on the first day as a part of the course introduction and then reference it throughout the semester. It serves as a handy tool for reminding students of "where we've been and where we're going" as we progress through the course material. In addition to the full syllabus, one author distributes a handout of just the graphic syllabus and asks her students to bring it to each class (students who organize their class material with a binder often place the graphic syllabus in the plastic cover on the outside of the binder). At beginning and end of each chapter, the instructor references the graphic syllabus to "review" where the course has been and "preview" where it is going.

One instructor has found the graphic syllabus to be especially useful in demonstrating concept mapping for students. After seeing the graphic syllabus as an example, students then make their own mind maps periodically throughout the semester in introductory microeconomics. In light of survey results, the instructor is considering having the students create their own mind maps, either individually or as a large group, of the course at the end of

the semester as an over-arching review and to help the instructor make any necessary changes to the existing graphics on the syllabus to enable future students to understand the course in a more clear manner.

In our principles courses, we have also used the graphic syllabus as a review exercise for exams. During the review, the graphic syllabus is displayed on the screen at the front of the classroom and students write down as much detail about each topic as they can remember without the aid of their notes. They also jot down questions they have. Then they reference their notes for those areas where they are "stuck" or where they have questions. After about 20 minutes of working this way, the remainder of the class is for question and answer. We have found this technique facilitates a very productive review session. Students are primed to ask questions on the topics on which they need the most instructor-assisted review.

In our previous experience with statistics courses, the progression from "descriptive statistics" to "probability" to "inferential statistics" has confused students. With the graphic syllabus, we have a concrete artifact to continually refer back to. This seems to have helped students to understand the linkages between the topics better and to have a clearer picture of the overarching goal of the course. We have also referenced the pie chart as an example of visual ways of representing data. It provides a tangible, real-world application that the students fully understand.

Anecdotally, students do seem to reference the extended graphic syllabus more often than they did our traditional syllabi. During class time, we have witnessed students (unprompted!) pull out the syllabus to look something up before asking a clarification on an upcoming due date or assignment. We have also overheard students telling other students where to find a particular piece of information, for example, "that policy's the one with the three boxes".

To assess students' perceptions of the syllabus, one instructor created two versions of her syllabus: an extended graphic syllabus and a traditional text syllabus for principles of microeconomics. The extended graphic syllabus presented the course information in a concept map format, and the assessment, or grading, section presented the grade distribution in a pie chart format. (The extended graphic syllabus also included graphics for the attendance policy and the learning objectives, but these sections were not the focus of the survey assessment). On the same page as the pie chart, each assessment type was described in its own box alongside the relevant "slice" of the pie chart.

On the first day of class, the instructor randomly passed out the two different versions to the students but did not tell the students that there were two versions. The students silently read through the syllabus on their own, and they were instructed to review the course description and assessment sections in particular. They then filled out a survey aimed at gauging the students' perceived understanding about the course and their interest in taking the course (Table 2). After completing the initial survey, the students were then provided with the other version of the syllabus and asked to read through this version silently. After reflecting upon both versions, the students filled out a second survey aimed at determining which version of the syllabus they believed was clearer and gave them a better understanding of the course and which version they preferred overall (Table 3). In addition to standard questions, there was also space for students to comment on why they chose their preferred version and to make any other comments. Finally, students were asked to turn in the version of the syllabus that they did not want to keep to refer back to throughout the semester.

Table 2 presents the results from the first survey. The survey was a Likert scale survey, with responses ranging from "Strongly Agree" to "Strongly Disagree." For each question, the top

row represents the responses from students who received the graphic syllabus, while the second row represents the responses from those who received the traditional syllabus. Fourteen students received the graphic syllabus first, while 11 received the traditional syllabus first. While there are no major differences between students' understanding of the course description, it appears that students may be more comfortable with the course description from the traditional syllabus. Those who received the graphic syllabus students were more likely to disagree with the statements that they could explain in plain language what the course was about, that the course description gave them a clear understanding of how the topics fit together in the bigger picture, and that they were excited to take the course compared to those who received the traditional syllabus.

Table 2
Student First-Impressions & Understanding of Syllabi

	Strongly Agree	Agree	Neither Agree nor Disagree	Disagree	Strongly Disagree	Syllabus Reviewed
Course Description						
After reading the course description, I understand what I'll be learning about this semester.	50%	50%	0%	0%	0%	Graphic
	55%	45%	0%	0%	0%	Traditional
After reading the course description, I can explain in plain language what this course is about.	14%	64%	14%	7%	0%	Graphic
	27%	55%	18%	0%	0%	Traditional
The course description gave me a clear understanding of how the topics fit together in the bigger picture.	54%	31%	8%	8%	0%	Graphic
	45%	36%	18%	0%	0%	Traditional
After reading the course description, I am excited to take this course.	43%	14%	21%	21%	0%	Graphic
	18%	55%	27%	0%	0%	Traditional
Assessments						
After reading through assessments, I understand how my grade will be determined.	100%	0%	0%	0%	0%	Graphic
	73%	27%	0%	0%	0%	Traditional
After reading through assessments, I understand the different components that will go into my grade.	93%	7%	0%	0%	0%	Graphic
	64%	36%	0%	0%	0%	Traditional
After reading through components, I understand what is expected of me this semester.	71%	29%	0%	0%	0%	Graphic
	73%	27%	0%	0%	0%	Traditional
I found it easy to understand how each component factors into my final grade.	79%	21%	0%	0%	0%	Graphic
	73%	27%	0%	0%	0%	Traditional

With respect to the assessment section of the syllabus, both syllabi contained information regarding the different assessment types and how much each type would count towards their final grade. The traditional syllabus listed the components along with their weighting, while the graphic syllabus presented the weighting in a pie chart. While all students, regardless of which version of the syllabus they received, either agreed or strongly agreed with all of the statements regarding the assessment section, in nearly all cases, those who received the graphic syllabus were slightly more likely to strongly agree with the statements. All of the students who received the graphic syllabus first strongly agreed that they understood how their grade would be determined, compared with 73% of those who received the traditional syllabus first. Ninety-three percent of the graphic syllabus students responded that they strongly agreed that they understood the different components that would go into their grade, compared with 64% of those who received the traditional syllabus.

The differences in the responses for the first survey could be because students at this point were unaware of how to read through a concept map as the instructor did not discuss the syllabi with them ahead of time. Also, students may not have had prior exposure to a concept map, and the familiarity of the traditional syllabus may comfort them. Similarly, students are much more familiar with pie charts, and the students who received the graphic syllabus appeared to find the assessment section of the syllabus clearer than those who received the traditional syllabus.

The second survey asked a variety of questions requiring the students to compare the two syllabi and choose which one they thought best answered the question. Table 3 presents the results. In 5 of the 9 questions, students preferred the graphic syllabus over the traditional. Fifty-six percent of students thought that the graphic syllabus gave them a better understanding of the big picture of the course, while 60% thought the traditional syllabus best explained what the course was about.

Oddly, while 60% of students thought the traditional syllabus was clearer in how their final grade would be determined, 72% thought the graphic syllabus more clearly explained the different components of their grade and 60% thought the graphic syllabus made it easier to understand the different components of their grade.

Interestingly, 84% of students thought the traditional syllabus was easier to understand, though 48% said they preferred the graphic syllabus overall and 40% preferred the traditional syllabus. Sixty percent of the students reported that the graphic syllabus made them more excited to take the course compared to only 32% who thought the traditional one made them more excited.

After the students filled out the surveys, the instructor explained the graphic syllabus to the students and reviewed the course expectations as is typical on the first day of class. At the end of the class, the instructor asked students to turn in the version of the survey they did not want to keep. Seventy-three percent of students kept the graphic syllabus over the traditional, indicating a revealed preference for the graphic syllabus.

Table 3
Student Preferences of Syllabi Type

Question	Graphic	Traditional	Neither
Which syllabus do you believe gave you a better understanding of the big picture of this course?	56%	44%	0%
Which syllabus best explains to you what the course is about?	40%	60%	0%
Which syllabus do you find easier to understand?	12%	84%	4%
Which syllabus do you think will be easier to refer to when you have questions about the course?	44%	44%	12%
Which syllabus makes it clearer to you how your grade will be determined?	40%	60%	0%
Which syllabus do you think clearly explains the different components going into your grade?	72%	16%	12%
Which syllabus makes it easier for you to understand the different components of your grade?	60%	36%	4%
Which syllabus makes you more excited to take this course?	60%	32%	8%
Which syllabus do you prefer overall?	48%	40%	12%

In sum, it appears that after students are lead through the graphic syllabus, especially if they have not had exposure to a graphic syllabus or if they are unfamiliar with a concept map, students generally prefer the graphic syllabus over the traditional syllabus. Students commented that the graphic syllabus was more eye-catching and interesting to look at and that it was easier to see the breakdown of grades and course topics. They liked that it was "quick and to the point" and "cut and dry". One student reported that he preferred the graphic syllabus because it was "not cramped with information and it organized information that is appealing while forming a better scheme of the course." A number of students who identified themselves as visual learners reported they liked the graphic syllabus better because they are visual learners.

Of those who preferred the traditional syllabus, a few students liked the details included in the traditional syllabus that were streamlined from the graphic syllabus. One student liked the ease of finding things in the graphic syllabus, but liked the details in the traditional syllabus. Another student commented that she thought sentences and paragraphs are easier to comprehend than graphs. One student commented that he originally thought the graphic syllabus was hard to follow, but after the instructor explained the graphic syllabus, he preferred the graphic version even though he indicated in the survey he preferred the traditional version.

The survey results indicate that the graphic syllabus may be better than the traditional syllabus in generating excitement for a course that many students are "forced" to take. It also clarifies the course for the visual learners, as theory predicted. Familiarity appears to influence students' initial impressions. Once students understand concept or mind mapping, most students prefer the graphic syllabus.

This survey method had some limitations. First, due to timing constraints, the entire sample was limited to 25 students in introductory microeconomics courses only. Second, the

instructor did not want to bias students, and so she did not take the time to walk the students through the graphic syllabus. For many students, this was the first time they had seen a concept map. Since the graphic syllabus is the instructor's interpretation of the course, explaining the syllabus, especially if it is the students' first exposure to a graphic syllabus, may help alleviate some confusion and overwhelm on the first day. Future work calls for more empirical investigations on students' experiences with the extended graphic syllabus. An assessment could provide students with multiple graphic syllabi examples to see which they prefer and to determine how the graphics on the syllabus influence the student opinion. Another option is administering instruments that capture qualitative data on attitudes and perceptions from students viewing either a traditional course outline or a graphic syllabus in different courses. Finally, a "syllabus quiz" at the midpoint in the semester, to a class of students working with a traditional syllabus and a class working with an extended graphic syllabus, could help determine if students retain more information. An additional experiment would measure the effectiveness of using the graphic syllabus as a review for an exam, versus other review techniques.

The two major drawbacks to the extended graphic syllabus are that it is not accessible to students with visual impairments (i.e. incompatible with screen readers) and it may run afoul of rules requiring a common syllabus across sections of the same course. In light of this and our survey results, we recommend that instructors keep a traditional "all text" syllabus to comply with rules set forth by the Americans with Disabilities Act. The instructor can post both versions of the document on the course learning management system and can distribute the extended graphic syllabus in class. Providing a traditional syllabus will also make the class more accessible for those students who do reveal a preference for the traditional, familiar format syllabus. For instructors teaching courses with common syllabi, we recommend creating an extended graphic syllabus as a supplement to the official course document. This supplement would include the instructor's contact information, a classic graphic syllabus, and the elements of the extended graphic syllabus that the instructor wants to draw the students' attention to, perhaps the course learning objectives, the course grading weights, and important due dates. Both documents could be distributed in class and posted on the course learning management system.

ACKNOWLEDGEMENT

Special thanks to participants of the 2011 Economics Teaching Conference (New Orleans, LA) and three anonymous referees for helpful comments.

REFERENCES

Allgood, S., W. Bosshardt, W. Van Der Klaauw, & M. Watts (2004). What students remember and say about college economics years later. *The American Economic Review,* 94(2), 259-265.

Ausubel, D.P. (1960). The use of advance organizers in the learning and retention of meaningful verbal material. *Journal of Educational Psychology,* 51, 267-272.

Ausubel, D.P. (1963). Cognitive structure and the facilitation of meaningful verbal learning. *Journal of Teacher Education,* 14(2), 217-222.

Baitz, I. (2009). Concept mapping in the online learning environment: a proven learning tool is transformed in a new environment. *International Journal of Learning,* 16(8), 285-291.

Biktimirov, E. & L. Nilson (2003). Mapping your course: designing a graphic syllabus for introductory finance. *Journal of Education for Business,* 78(6), 308-312.

Budd, J. W. (2004). Mind maps as classroom exercises. *Journal of Economic Education,* 35(1), 35.

Carnegie Mellon University (n.d.) In Teaching Excellence & Educational Innovation website. The syllabus. Retrieved November 13, 2011 from http://www.cmu.edu/teaching/designteach/design/syllabus/

Clark, J. & L. Paivio (1991). Dual coding theory and education. *Educational Psychology Review*,3(3), 149-210.

Harnish, R.J. & K.R. Bridges (2011). Effect of syllabus tone: students' perceptions of instructor and course. *Social Psychology of Education: An International Journal*, 14(3), 319-330.

Horton, P.B., A.A. McConney, M. Gallo, A.L. Woods, G.J. Senn, D. Hamelin, et al. (1993). An investigation of the effectiveness of concept mapping as an instructional tool. *Science Education*, 77(1), 95-111.

Manoli, P. & M. Papadopoulou (2012). Graphic organizers as a reading strategy: research findings and issues. *Creative Education*, 3(3), 348-356.

Marangos, J. (2003). The effectiveness of concept maps in introductory microeconomics. *Economic Papers*, 22(4), 74-82.

Marangos, J. & S. Alley (2007). Effectiveness of concept maps in economics: evidence from Australia and USA. *Learning and Individual Differences*, 17(2), 193-199.

Nilson, L.B. (2007). *The graphic syllabus and the outcomes map: communicating your course (First Edition)*. San Francisco, CA: Jossey-Bass.

Nilson, L.B. (2010). *Teaching at its best: a research-based resource for college instructors.* San Franciso, CA: Jossey-Bass.

Novak, J.D. & A.J. Canas (2008). The theory underlying concept maps and how to construct and use them. Technical report IHMC CmapTools (2006-01; Rev 01-2008). Florida Institute for Human and Machine Cognition. http://cmap.ihmc.us/Publications/ResearchPapers/TheoryUnderlyingConceptMaps.pdf.

Vark-Learn. (2011, November). Table five: groups and the percentage of V, A, R and K options chosen. Retrieved April 20, 2012 from http://www.vark-learn.com/english/page.asp?p=research.

Vekiri, I. (2002). What is the value of graphical displays in learning? *Educational Psychology Review*, 14 (3), 261-312.

Willerman, M. & R.A. Macharg (1991). The Concept Map as an Advance Organizer. *Journal of Research in Science Teaching*, 28(8), 705-711.

Wills, S. & E. Ellis (2010). The theoretical and empirical basis for graphic organizer instruction. Retrieved June 29, 2012 from MakesSenseStrategies.com.

ARE INCOME DISTRIBUTION PROGRAMS DESIGNED TO REDISTRIBUTE INCOME? DO THEY? EVIDENCE FROM THE STATES

John J. Bethune, Barton College

ABSTRACT

Income inequality has generally been viewed as a problem and concern for policy makers. Income redistribution programs are a means to address this problem. These programs typically consist of progressive taxation to lower the incomes of the highest earners and transfer programs designed to raise the incomes of the lowest earners.

The Gini Coefficient is the most popular tool to measure income inequality within nations or states. It measures the degree of inequality between the quintiles of a region's population. This paper calculates the Gini Coefficient for each of the 50 states and the District of Columbia before any redistribution programs are implemented. It then recalculates the Gini Coefficients after taking into account redistribution efforts at the federal and state level. A comparative analysis is then conducted to determine how much more equality results from these programs from an absolute and a relative perspective.

It might be expected that states with higher measures of income inequality would have more aggressive income redistribution programs. This paper examines the extent to which this expectation is accurate. Other policies that are promoted to remedy the problem of income inequality are higher minimum wages and greater spending on education. The effectiveness of these policies is also addressed.

INTRODUCTION

Income inequality has generally been viewed as a problem and a concern for policy makers. Income redistribution programs are a means to address this problem. These programs typically consist of progressive taxation to lower the incomes of the highest earners and transfer programs designed to raise the incomes of the lowest earners.

The Gini Coefficient is the most popular tool to measure income inequality within nations or states. It measures the degree of inequality between the quintiles of a region's population. This paper calculates the Gini Coefficient for each of the 50 states and the District of Columbia before any redistribution programs are implemented. It then recalculates the Gini Coefficients after taking into account redistribution efforts at the federal and state level. A comparative analysis is then conducted to determine how much more equality results from these programs from an absolute and a relative perspective.

It might be expected that states with higher measures of income inequality would have more aggressive income redistribution programs. This paper examines the extent to which this expectation is accurate. Other policies that are promoted to remedy the problem of income inequality are higher minimum wages and greater spending on education. The effectiveness of these policies is also addressed.

Prior Research

The Gini Coefficient is a statistic developed by an Italian statistician named Corado Gini that measures the degree of income inequality in a region by quintiles. The value of the coefficient can range from 0, which would imply complete income equality, to 1, indicating that one person (or household) has all of the income. It is the most popular measure of the degree of income inequality in both nations and states.

While there have been various comparative studies of income distributions among nations, much less has been written about inequality at the state level in the U. S. One study (Econscious) used regression analysis in an attempt to identify the causes of inequality between the states. Explanatory variables tested included whether the state was a "right to work" state (no relationship), state income tax rates (higher state tax rates correlate with higher degrees of inequality), and population density (strongest correlation, with an R-square of 0.14). The information presented is a descriptive analysis and no policy implications are offered.

Another article (Leigh) addresses the question of whether redistributive state tax policies reduce inequalities. Using data from 1977 to 2002, the author concludes that redistributive taxation reduces income inequality and that these policies do not slow the growth of per capita income.

In a 2012 publication (Cooper, et. al.) the authors analyze the effects of federal and state taxes on the distribution of income within individual states. The resulting comparison of pre and post tax Gini Coefficients show substantial reductions in inequality in all states. They find that most of the gains result from federal tax policy, though there is substantial variation among the states.

These prior studies have addressed issues in what might explain differences in income distributions between the states, and how effective redistributive policies have been in reducing inequality. This study will try to determine if state policies are actually designed to reduce inequality and if they are effective in improving a state's relative position in income inequality rankings.

Gini Coefficients

The 2010 Census provides the most comprehensive source of data for the questions we want to address in this study. Therefore, unless otherwise noted, all data sources are from 2010. Table B19081 (Census) contains the mean income for all households in each quintile by state. Using these figures, I calculated a Gini Coefficient for each state. These are included in the column labeled Own Gini in Table I. The second column, Popular Gini, contains the most widely

circulated measures of Gini Coefficients for the states, found in periodicals and posted on Wikipedia (Wikipedia).

The coefficients in the Popular measure tend to be higher than those in the Own calculation, however the correlation coefficient between the two measures is .949 with a P-Value of .000, so the measures track almost perfectly together.

The third column of Table I are the Gini Coefficients after adjusting for state and federal taxes and federal transfers (After Tx & Tr Gini). I calculated these by using data from the Congressional Budget Office (for Federal taxes by quintile), the Tax Foundation (for state tax rates) and the Census Bureau (Table DP03, for cash transfers and SNAP benefits). While there are some cash assistance programs from the states and noncash transfer programs to assist households in the lower quintile, these are difficult to quantify and impute an accurate value.

The last column in Table I contains the Gini Coefficients calculated by Cooper, et. al. in their 2012 study (After Taxes Gini). It accounts for all taxes including sales and motor fuels taxes. This calculation generally results in the lowest values for all Gini Coefficients presented here. The average value for each of the columns is:

Popular Gini	.453
Own Gini	.420
After Tx & Tr Gini	.376
After Taxes Gini	.314

By comparing the averages of the first two Ginis listed with the latter two, it is clear that tax and transfer policies do substantially reduce the absolute level of income inequality in the U.S. This would confirm the results from prior research that found this to be true. But do these policies change state rankings in any significant way?

Table I

State	Own Gini	Popular Gini	After Tx & Tr Gini	After Taxes Gini
Alabama	0.440	0.4720	0.392	0.325
Alaska	0.392	0.4220	0.356	0.311
Arizona	0.432	0.4550	0.380	0.331
Arkansas	0.432	0.4580	0.388	0.309
California	0.448	0.4710	0.392	0.330
Colorado	0.428	0.4570	0.388	0.327
Connecticut	0.428	0.4860	0.380	0.320
DC	0.492	0.5320	0.444	0.314
Delaware	0.408	0.4400	0.356	0.306
Florida	0.440	0.4740	0.400	0.341
Georgia	0.428	0.4680	0.392	0.320
Hawaii	0.396	0.4330	0.341	0.315
Idaho	0.396	0.4330	0.356	0.305
Illinois	0.432	0.4650	0.392	0.322
Indiana	0.412	0.4400	0.364	0.308
Iowa	0.396	0.4270	0.340	0.290
Kansas	0.412	0.4450	0.360	0.314
Kentucky	0.432	0.4660	0.392	0.311
Louisiana	0.444	0.4750	0.392	0.329
Maine	0.404	0.4370	0.356	0.285
Maryland	0.412	0.4430	0.372	0.322
Massachusetts	0.440	0.4750	0.392	0.324
Michigan	0.416	0.4510	0.389	0.312
Minnesota	0.408	0.4400	0.356	0.308
Mississippi	0.440	0.4680	0.392	0.331
Missouri	0.416	0.4550	0.380	0.308
Montana	0.412	0.4350	0.356	0.293
Nebraska	0.396	0.4320	0.356	0.300
Nevada	0.412	0.4480	0.380	0.321
New Hampshire	0.396	0.4250	0.356	0.322
New Jersey	0.432	0.4640	0.364	0.335
New Mexico	0.432	0.4640	0.392	0.318
New York	0.455	0.4990	0.408	0.322
North Carolina	0.432	0.4640	0.392	0.321
North Dakota	0.396	0.4330	0.356	0.305
Ohio	0.416	0.4330	0.372	0.301
Oklahoma	0.416	0.4520	0.365	0.312
Oregon	0.412	0.4540	0.360	0.301
Pennsylvania	0.432	0.4490	0.392	0.312
Rhode Island	0.432	0.4610	0.388	0.309
South Carolina	0.428	0.4610	0.384	0.309
South Dakota	0.412	0.4420	0.372	0.313
Tennessee	0.432	0.4680	0.396	0.328
Texas	0.432	0.4690	0.396	0.347
Utah	0.392	0.4190	0.344	0.294
Vermont	0.408	0.4440	0.356	0.296
Virginia	0.428	0.4590	0.388	0.329
Washington	0.408	0.4410	0.372	0.325
West Virginia	0.412	0.4541	0.368	0.301
Wisconsin	0.396	0.4300	0.352	0.295
Wyoming	0.392	0.4230	0.344	0.303

For the purpose of this analysis it is important to consider the relative ranks of the states for each measure of the Gini Coefficients presented in Table I. These rankings are included in Table II. A ranking of 1 indicates the state with the most equal distribution of income and a ranking of 51 would be the state with the most unequal income distribution

Table II

State	Popular Gini Rank	Own Gini Rank	After T & T Rank	After Taxes Rank
Alabama	45	45.5	41.0	40.5
Alaska	2	2.0	10.5	21.5
Arizona	27	38.0	27.5	47.5
Arkansas	30	38.0	32.5	19.0
California	44	49.0	41.0	46.0
Colorado	29	30.0	32.5	42.0
Connecticut	49	30.0	27.5	31.5
DC	51	51.0	51.0	27.5
Delaware	13	13.5	10.5	14.0
Florida	46	45.5	49.0	50.0
Georgia	40	30.0	41.0	31.5
Hawaii	8	7.0	2.0	29.0
Idaho	8	7.0	10.5	12.5
Illinois	37	38.0	41.0	36.5
Indiana	13	19.5	18.5	16.0
Iowa	5	7.0	1.0	2.0
Kansas	20	19.5	16.5	27.5
Kentucky	38	38.0	41.0	21.5
Louisiana	47	48.0	41.0	44.5
Maine	12	11.0	10.5	1.0
Maryland	18	19.5	23.5	36.5
Massachusetts	47	45.5	41.0	39.0
Michigan	23	25.5	35.0	24.0
Minnesota	13	13.5	10.5	16.0
Mississippi	40	45.5	41.0	47.5
Missouri	27	25.5	27.5	16.0
Montana	11	19.5	10.5	3.0
Nebraska	7	7.0	10.5	7.0
Nevada	21	19.5	27.5	33.5
New Hampshire	4	7.0	10.5	36.5
New Jersey	34	38.0	18.5	49.0
New Mexico	34	38.0	41.0	30.0
New York	50	50.0	50.0	36.5
North Carolina	34	38.0	41.0	33.5
North Dakota	8	7.0	10.5	12.5
Ohio	8	25.5	23.5	9.0
Oklahoma	25	25.5	20.0	24.0
Oregon	26	19.5	16.5	9.0
Pennsylvania	22	38.0	41.0	24.0
Rhode Island	32	38.0	32.5	19.0
South Carolina	32	30.0	30.0	19.0
South Dakota	17	19.5	23.5	26.0
Tennessee	40	38.0	47.5	43.0
Texas	43	38.0	47.5	51.0
Utah	1	2.0	3.5	4.0
Vermont	19	13.5	10.5	6.0
Virginia	31	30.0	32.5	44.5
Washington	16	13.5	23.5	40.5
West Virginia	23	19.5	21.0	9.0
Wisconsin	6	7.0	5.0	5.0
Wyoming	3	2.0	3.5	11.0

Income Inequality And State Policy

There are primarily two types of redistribution activities that states might engage in to address problems of income inequality. These are taxes and transfer programs. We can measure the extent to which states are using taxes to redistribute income by examining the top marginal income tax rates in each state. The Tax Foundation has this information for 2010.

It would be expected that if states in general are using their income taxes to address the problem of income inequality then states with higher marginal tax rates would correlate with states that have higher pre-tax Gini Coefficients. This is not evident from an analysis of simple correlation coefficients.

If we correlate top state marginal tax rates with the state Gini Coefficients we get values of .127 for Own Gini (P-Value = .375) and .162 for Popular Gini (P-Value = .257) indicating no significant relationship between these measures of income inequality and top marginal tax rates in the states.

Similarly, if we correlate the state rankings based on top marginal tax rates and state rankings for the Gini Coefficients we get even weaker results of .011 for the Own Gini Rank (P-Value = .938) and .043 for the Popular Gini Rank (P-Value = .764).

Even though states do not appear to be using marginal tax rates to address the problem of income inequality, do the policies have some effect on the distribution of income? Correlating the top marginal rates with the after tax measures of income inequality we get an After Tx & Tr Gini coefficient of -.101 (P-Value = .480) and an After Taxes coefficient of -.345 (P-Value = .013). For these relationships the negative signs indicate that the higher marginal tax rates are associated with lower Gini calculations, and there is some significance in the After Taxes correlation coefficient, as indicated by the lower P-Value.

Comparing the marginal tax rate rank with the After Tx & Tr Rank we get a coefficient of .011 (P-Value = .938) and for the After Taxes Rank the coefficient is -.347 (P-Value = .013). This would again reinforce the implication that higher marginal tax rates are associated with lower measures of income inequality with respect to the After Taxes Gini measure. This would imply that federal and state tax policies, when broadly viewed, do achieve lower degrees of income inequality.

Measuring the total value of transfers to lower income households is a fairly difficult figure to quantify. A study of the total level of welfare benefits by the states was conducted by Tanner and Hughes and this provides the best measure for the question being address here: Are transfer programs designed to address disparities in income at the state level?

The total value of both federal and state benefits are included in their analysis and that can be correlated with the Gini Coefficients to determine if these transfer programs are targeted at reducing income inequality in the states and how well they are doing in achieving this goal (intended or unintended).

The total value of transfers for each state has no significant correlation with the Popular Gini (-.111, P-Value = .438) and actually has a negative sign as does the correlation with the Own Gini (-.156, P-Value = .274). This would indicate that higher Gini Coefficients correspond to lower total benefits, which would be opposite of what might be expected.

Similarly, the rankings of states welfare benefits have no significant correlation with the rankings of the states by Popular Gini (-.082, P-Value = .569) and Own Gini (-.036, P-Value = .804). In this case however, the negative signs would be expected because the higher ranked states in benefits (most benefits = 1) should negatively correlate with the more unequal distribution of income (most unequal = 51).

Given these relationships we can conclude that state and federal transfer payments are not systematically intended to address the unequal distribution of incomes between states.

The next question is, do the transfer payments alter income distributions between the states after taxes and transfers have been reflected in the Gini calculations. Correlating the total benefits with the After Taxes Gini shows no relationship (.008, P-Value = .957) and the correlation with the After Tx & Tr Gini actually yields a negative sign (-.231, P-Value = .103) implying that higher benefits are associated with a more equal income distribution.

The state rankings by transfer payments also do not correlate with the rankings by After Taxes Gini (-.023, P-Value = .872) or After Tx & Tr Gini (.060, P-Value = .676).

All of the above would indicate that relative state benefits are not intended to reduce relative income inequality and they do not succeed in reducing relative income inequality.

Recent Policy Proposals To Reduce Income Inequality

A policy proposal to increase the federal minimum wage has been offered as a mechanism for reducing income inequality. If this was an effective approach then we would expect states with higher minimum wage laws to have lower Gini coefficients then states that only follow federal minimum wage policies. However, this does not appear to be the case.

20 states have mandated higher minimum wages than the federal rate. These states have an average Popular Gini score of .462. States without higher minimum wages have an average of .448. Similarly, the Own Gini average for the higher minimum states is .428 and the average for the other states is .415.

State Gini rankings also do not support the notion that higher minimum wages reduce income inequality. States with higher mandated minimum wages have an average ranking of 30.35 with Popular Gini and 31.05 with Own Gini (recall that a ranking of 51 implies the most unequal distribution of income). States without the higher minimum wage mandates have an average rank 21.68 with Popular Gini and 22.74 with Own Gini.

States with higher minimum wage laws rank higher in income inequality measures (where a higher rank indicates a more unequal distribution of income) than states without such laws.

Greater spending on education has been proposed as a means to lower income inequality. If this is an effective measure we would expect states that spend more per pupil would have lower levels of income inequality than states that spend less.

Correlating the amount spent by states per pupil with Popular Gini we get a weak but significant relationship (.227, P-Value = .049) indicating that states that spend more on education have a higher level of income inequality. The correlation with Own Gini (.211, P-Value = .137) yields the same sign but no significance.

If we consider state rankings there appears to be no significant correlation with education spending ranked by state with Popular Gini rankings (-.089, P-Value = .536) or Own Gini rankings (-.121, P-Value = .396).

States that spend more on education do not achieve a more equal distribution of income.

Some have argued that economic growth is the most important factor in reducing income inequality. If we examine the growth in GDP of the states from 2000 to 2010 and their relationship to the Gini Coefficients in 2010 we get correlated values of -.269 (P-value = .056) for Popular Gini and -.292 (P-value = .038) for Own Gini. These show that income inequality is negatively correlated with economic growth and they are significant at the 10 percent level for Popular Gini and at the 5 percent level for Own Gini.

The same relationships exist if we compare the state rankings for GDP growth through 2010 with the Popular Gini (-.269, P-value = .067) and Own Gini (-.284, P-value = .043) state rankings. These values suggest there is some association between stronger economic growth and less income inequality.

If we compare state GDP growth from 2010 through 2012 we find that the correlation with Popular Gini (-.087, P-value = .542) and Own Gini (-.121, P-value = .397) are still negative, but insignificant for this time frame. The same holds true if we compare state rankings of GDP growth and Popular Gini (-.116, P-value = .419) and Own Gini (-.138, P-value = .333) state rankings. While economic growth rates from 2010 onward remain negatively associated with income inequality there is not a significant relationship during this limited time frame.

SUMMARY

State income tax policies do not appear to be designed to address the problem of income inequality and they do not succeed in reducing this inequality. When a broader range of taxes are considered there is a significant correlation with relative inequality reduction at the state level.

Transfer programs implemented at the federal and state level also do not appear to be designed to reduce relative income inequality between the states and they do not result in less relative inequality.

Higher minimum wage laws are associated with greater, not less, income inequality and would not appear to be effective measures for achieving a more equal distribution of income. Similarly, higher levels of spending on education do not achieve a more equal distribution of income.

Higher rates of economic growth are significantly associated with more equal distributions of income, which might imply that, for an individual state, policies to promote economic growth are the key for achieving less income inequality relative to other states.

REFERENCES

Census, 2010 American Communities Survey, United States Census Bureau, Department of Commerce, http://factfinder2.census.gov/faces/nav/jsf/pages/index.xhtml

Cooper, Daniel H., Lutz, Byron F. and Palumbo, Michael G., "Quantifying the Role of Federal and State Taxes in Mitigating Wage Inequality, Finance and Economics Discussion Series Divisions of Research & Statistics and Monetary Affairs, Federal Reserve Board, Washington, D.C. January, 2012

Congressional Budget Office, The Distribution of Household Income and Federal Taxes, 2009. http://www.cbo.gov/publication/43373

Econscious, "Which State has the Most Income Inequality? Texas or New York? Why?" http://econscius.wordpress.com/tag/gini-coefficient/

Leigh, Andrew, "Do Redistributive State Taxes Reduce Inequality?" *National Tax Journal*, Vol. LXI, No. 1, March, 2008.

Tanner, Michael and Hughes, Charles, The Work Versus Welfare Trade-Off: 2013, The CATO Institute, Washington, DC, 2013.

The Tax Foundation, http://taxfoundation.org/

Wikipedia, "List of U. S. States by Gini Coefficients," http://en.wikipedia.org/wiki/List_of_U.S._states_by_Gini_coefficient

COLLABORATIVE PROBLEM SOLVING PROMOTES STUDENTS' INTEREST

Khoo Yin Yin, Sultan Idris Education University, Malaysia

ABSTRACT

The aim of the study was to investigate the effects of collaborative problem solving (CPS) on students' interest of learning economics. The economics concept was a difficult subject for pre-university students to comprehend. A quasi-experimental method was applied in this study where 294 students were selected randomly from ten schools and divided into three groups (CPS1, CPS2 and CG). Surveys and interviews had been carried out before and after intervention. CPS1 outperformed in interest compared to CPS2 and CG. CPS2 also performed quite well in comparison to CG. The findings of qualitative responses also showed positive results in CPS. CPS can be applied in schools not only in economics but also other subjects.

Field of Research: collaborative problem solving, Economics Education

INTRODUCTION

Economics teaching in the pre-university level seem important as a large number of pre-university students are partial to take the subject on. This phenomenon was supported by the report from the Malaysian Examination Council (2008) which showed that there were 30,737, 29103 and 23,570 candidates who took economics in 2005, 2006 and 2007 respectively.

Although there were a great number of students who undertook economics yearly, the overall achievement in this subject was declining from year to year. This fact can be proven by the pre-university public examination results whereby a mere 52.33%, 51.84% and 49.30% passing rate was obtained in the years of 2005, 2006 and 2007 respectively (Malaysia Examination Council, 2008). Consequently, most students lost interest in this subject (Khoo Yin Yin, 2008).

One of the main factors of lack of performance in this subject was due to art students who undertook economics that were not keen in analyzing and understanding the concept while applying the element of mathematics. Johnston, James, Lye and McDonald (2000) suggested the best way of teaching economics is the implementation of active learning. In spite of this, the most popular teaching method amongst teachers is the "chalk and talk" method (Becker & Watts, 2001). This is a traditional teaching method where students are not required to be learning actively. Prior research shows that active learning promote students' learning interest in various subjects (Leung Yin Bing & Hui, 2009; Matveev & Milter, 2010; Chen Yuqing, Peng Xiaoshan & Sun Jian, 2010). However, there is a research gap between the implementation of CPS and pre-university economics students. Prior reviews from Jane and Jiri (2009), gave an overview of the effect of the implementation of active learning in secondary schools.

Previous studies have shown significant results between active learning and students' interest (Goldman, Cohen & Sheahan, 2008; Muhammad Akhlaq, Mukhtar Ahmad Chudhary, Samina Malik, Saeed-ul-Hassan & Khalid Mehmood, 2010). Therefore, CPS is claimed to be a practical strategy for fostering students' interest and performance. In addition, CPS is a theory and concept of active learning with the element of practical application on economics. This study will measure the effect of CPS on improving students' interest on economic learning.

CONCEPTUAL FRAMEWORK

The conceptual framework of this study is based on Adult Learning Theory (Knowles, Holton and Swanson, 1998) and Vygotsky's zone of proximal development (ZPD) theory (1997). According to Knowles, Holton and Swanson (1998), adults were autonomous and self- directed. Their teachers might involve adult participants in their active learning process and serve as facilitators. ZPD is the difference between what a learner can do without help and what he or she can do with help under the guidance of a facilitator. ZPD is easier to be achieved during activities, such as in a collaborative group (Vygotsky, 1997).

In this study, students are exposed to new information during discussions with peers. They had to resolve idea between prior understanding of old information and new information provided by peers. CPS group members had to share their ideas and helped their peers to achieve the ZPD learning zone by a more capable peer and teacher (Vygotsky, 1997). Teachers played an important role as facilitators in assisting and explaining to students. On the other hand, social interaction among group members is crucial in students' acquisition of new knowledge and critical thinking skills (Vygotsky, 1997) as well as promoting students' interest.

LITERATURE REVIEW

Collaborative Problem Solving

According to Wehmeier (2000) in the Oxford Advanced Learner's Dictionary, problem solving is defined as the action of finding a way to deal with a problem. Wehmeier (2000) also describes the term collaboration as the involvement of a few people in a group who work together. Both problem solving and collaboration signify that students are to present a question and the group members are to solve the problem collaboratively (together). Students are expected to use resources to find knowledge based on self-directed learning. The success of the problem solving depends on the evaluation by the problem solvers, self-directed students and the cooperation of group members (Barrows, 1994). Teachers need to assist students during the CPS process (Barrows, 1994). Teachers act as the expert in the problem solving process but they must be trained in this field because the success or failure of the problem solving depends highly on the teachers' function. Basically, the transformation of problem solving depends on the analysis of the capability of the students as well as the planning of the teacher during the process.

In point of fact, the CPS method originated from the Problem Based Learning method used in 1960 in a medical school in the United States of America. The CPS method is most suitable for heuristic tasks. Heuristic tasks include a complex question that needs great knowledge and high thinking skills to solve. Currently this method is used widely in the tertiary

level for fields such as computers, engineering, business and architecture. CPS is also used in the subject of economics in high schools and universities across Australia.

Among the advantages of CPS include building problem solving skills in addition to increasing self-directed learning and lifelong learning. It also promotes a good relationship between group members, increases intrinsic motivation to learn and enhances the interaction between students and the facilitator.

Blumberg and Michael (1992) stated that students in problem based learning (PBL) possess the action and reaction which displayed self-directed learning skills. Mauffette, Kandlbinder and Soucisse (2004) supported their views. Thorley and Gregory (1994) believed that students who worked together in a group or collaboratively generated knowledge through discussion and this advantage outweighs the individual learning. Members who collaborated will have a clearer understanding of concepts or theories taught through the exchange of ideas and produced work of better quality.

According to Baurer (2003), the study of CPS is to ensure that the group functions more effectively. The teacher should ensure that the group is dynamic and proceed smoothly. Problem solving method emphasises on the students' attendance, planning, involvement, equal contribution to the group project and motivates learning. These findings are supported by Gokhale (1995) who advocates that collaborative learning can develop critical thinking through discussions, classification and evaluation of other people's ideas.

Besides that, CPS enhances metacognitive skills. Bonk and Cunningham (1998) posits that learning will happen quicker when students process self-directed skill before discussion. Problem solving learning is the learning and teaching process which develops students' ability based on the metacognitive level. The success of the CPS is not based solely on knowledge but is based on the problem solving method to achieve targets (Gijselaers, 1996).

The types of different learning for CPS include conceptual development and cognitive strategies. The formation of conceptual understanding includes the development of new knowledge schema and the assimilation of schema content. Cognitive strategies include critical thinking skills, strategy learning and metacognitive skills (Nelson, 1999).

Collaborative problem solving enables the advancement of critical thinking in the economics subject through assignments that reflect the real world situations, integration of learning activities which incorporates subject knowledge, types of thinking, skills and available opportunities which enable students to use more writing with pencil and paper to enhance motivation in providing meaningful tasks to students. Besides that, frame work may be used for active and collaborative learning which enable students to solve problems through analysis, application and resources acquired.

In the studies conducted by Johnston et al. (2000), Brooks and Khandler (2002), it is found that students who learn in a collaborative group obtained better results in examinations. Johnston et al. (2000) studied the CPS among year two students in the University of Melbourne and found out that it is one of the active learning methods that could stimulate learning. The project carried out could enhance communication skills and group work skills which increased learning. Collaborative learning assists students' discussion and integrates new ideas to learning in depth. Based on the findings, students' academic achievement after undergoing a nine-month project revealed that there were positive relationships between projects which utilised the CPS method. Students in these projects used a longer duration of time to prepare questions before proceeding to tutorial classes. Students' attendance showed an increase of 3% compared to the traditional tutorial classes. The findings of this study showed a change in the respondents' score

which is not equivalent to the two months duration used for preparation. Students' achievement was significant for the international students group but there was no significance for the Asian students group because they believed that the Asian students possess different learning styles. The findings from the study conducted by Johnston et al. (2000) also showed that a different duration is taken by the semester one experimental group while the control group was taken during the second semester. A different duration of time is not suitable to conduct the experimental study as students had experienced psychology and mental changes after one semester.

According to Alexander and McDougall (2001), tutors and students showed significance in the change of the traditional tutorial method to the CPS method as a new tutorial method. Mergendoller, Lahart and Mass (2002) studied the difference between the CPS method and the traditional teaching in a secondary school. From the findings, students who participated in the research programme did not show any significance in the change of the students' attitude towards economics and interest in the CPS method. This finding showed the opposite situation from that of which was hypothesised based on the review of the medical education work; CPS method is more effective compared to conventional learning. Based on gender, the raw data collected showed that female students obtained more benefits from the CPS method compared to male students even though there was no significance at the .05 level. Students involved in the CPS method showed progress in knowledge compared to the conventional method in the raw data. The main factor of the CPS method is that some students were involved in the problem solving whereas other students did not get involved but was merely waiting for the answers. Usually only one or two students worked while the rest became free passengers in the group.

Findings from studies are mostly taken from university students. However, the study conducted by Mergendoller, Maxwell and Bellisimo (2002) was carried out in a secondary school. The findings from the study conducted by Khoo Yin Yin and Zakaria Kassim (2005) on 200 form six economics students from four Penang secondary schools in 2004 showed that students studied economics through memorisation. They obtained higher scores in examinations if the examination questions were similar to their notes and work books. Students obtained better scores for lower thinking questions which were questions that tested knowledge, comprehension and simple application. On the other hand, for questions that were of high thinking levels such as synthesis and evaluation, students faced great difficulties. Amal Al-Dujarly and Hokyoung Ryu (2007) stated that CPS could develop a positive learning style. The findings showed that CPS is able to train students to develop an independent learning style.

All the methodology of the studies conducted was experimental and quantitative in the form of data analysis. One of the setbacks of the previous studies is that the experimental groups were collaborative in nature compared to individual learning. This setback will be given attention and will not be repeated by the researcher.

Theory Of Collaborative Problem Solving Method

CPS method is an active learning method. During the problem solving activity, students work together to ensure that the problem posed will be solved. The collaborative group is able to solve the problem posed and is able to identify its own weaknesses (Vygotsky, 1997). Problems prepared by teachers were solved by working with partners (Chiu Ming Ming, 2000). Collaborative problem solving method requires that a problem is solved in groups and not merely by an individual's ability. This is caused by lack of experience, different individual perspectives

and knowledge and experience levels about a different thing. There are many theories involved in the collaborative problem solving method such as instructional theory, behaviourist theory and constructivist theory. However, the focus of this study leads to the constructivist theory because this theory focuses on the mental activities which greatly influenced the learning outcomes through the collaborative problem solving method. The constructivist theory conflicted with the behaviourist theory which stressed on mental activities, knowledge originality and the way students develop knowledge from their actions. The collaborative problem solving method is learning based on the integration of cognitive and social perspectives to construct learning. Mergel (1998), considered the constructivist as a theory that involves learning from experience. According to Sally Hong (2002), constructivism is a theory for learning and philosophy for understanding. Learning is a constructive process in which economics students build knowledge based on prior knowledge. This matter is a process which fills the students' minds with information. This learning only allows economics students to retain facts or concepts in their memory and retrieve them when needed. The modern cognitive psychology states learning as a retaining process and based on concepts. New information can be used to collect and solve problems. Economics students are more suitable to be in collaborative learning which is a small group working together to solve problems.

METHODOLOGY

This study employed quasi experimental design with random assignment of 294 pre-university economics students from ten secondary schools in Penang, Malaysia. In order to control the threat of validity, all teachers had to teach the same microeconomics content using CPS method. Besides, teachers involved were given two sections of briefing prior to implementation of CPS method in the class to ensure proper usage of the method. The head of department was assigned to supervise teachers for the progress of the implementation of the CPS method .

This study was carried out in the middle of the year. The experiment took ten weeks to complete. The students in the selected schools were assigned randomly into three groups. There were CPS1, CPS2 and CG. CPS1 was a group that student learnt through CPS with fix working steps, whereas, CPS2 was a group that students learnt through CPS with free working steps. CG was a conventional group which function as a control group.

Pilot test was conducted to test for validity and reliability of the instruments. A set of questionnaires consisted of 10 items has found valid with reference to two lecturer in economics education. The overall Cronbach's Alpha reliability coefficients of questionnaire was obtained in .830. The pre and post questionnaires were given to 294 students who took part in this study.

Pre and post interview for 6 teachers and eight students was carried out. Their aim of the pre-interviews was to find out students' interest before implementation of CPS. The post-interview was to get their feedback about students' interest after implementation of CPS. The interview were tape-recorded and lasted approximately twenty minutes each participant. After the interviews, each word was coded, read in several times and analyzed using content analysis technique (Berg,1998).

FINDINGS

A) Analysis Of Questionnaire

Table 1 listed on mean of pre and post experiment CPS. One of the strongest benefits of CPS1 was built up confident of students to solve problem in economics (item 1). Item 1 also showed the highest difference of mean score. Item 2 and item 7 showed that the impact of CPS could help students' competence problem solving and understanding the microeconomics concept. It also indicated majority of the students liked group learning activities, it stated in item 6 and 10. In addition, CPS1 learning method showed the increased of mean on each item.

Table 1: Mean of pre-university students interest before and after implementing collaborative problem solving

Item	Question	Pre Experiment (CPS1)	Pre Experiment (CPS2)	Pre Experiment (CG)	Post Experiment (CPS1)	Post Experiment (CPS2)	Post Experiment (CG)
1	Confident with problem solving	3.342	3.663	3.659	4.026	3.794	3.590
2	Like economics theory	3.386	3.402	3.580	3.763	3.652	3.500
3	Interest in economics	3.316	3.424	3.455	3.544	3.652	3.523
4	Enjoy learning economics	3.658	3.739	3.839	4.044	3.674	3.681
5	Discuss economics topic	3.535	3.380	3.494	3.921	3.576	3.398
6	Discuss with peers during CPS	3.325	3.337	3.318	3.693	3.489	3.489
7	Problem solving become very simple	3.228	3.294	3.193	3.781	3.478	3.477
8	Present better task	3.544	3.359	3.330	3.798	3.587	3.625
9	More time to spend on this subject	3.500	3.380	3.330	3.816	3.489	3.511
10	study economics in group	3.491	3.391	3.490	3.895	3.717	3.625

Surprisingly, item 4 showed decreasing of enjoy learning economics after implementing CPS2. However, researchers found out the influence on CPS2 towards interest through item 12,3,5,6, 7,8 and 10. These items obtained high mean score which above 3.4. This meant CPS had influenced on fostering students interest significantly. On the other hand, students in CG didn't show the decreasing of mean on item 2, 4 and 5. The overall mean score for CPS1 was 3.828, CPS2 was 3.611 and CG was 3.542. These results indicated that CPS1 outperformed compare to CPS2 and CG.

B) Pre-Intervention Interview

Six teachers and eight students were interviewed before implementation of CPS. Teachers were asked about problem they were faced when teaching economics. Besides students were asked about their interest to economics subject.

i) Teachers' perspective

Most of the teachers said that their students did not study economics before when the were in secondary school. Therefore, they dislike this subject and found it difficult. The two quotes below revealed similar beliefs.

Paragraph 1

PIT5: "...students' attitude. Majority of them not interest, they feel economic is a difficult subject, especially those poor in Mathematic."
PIT3: " They didn't study, refuse to do exercise. Most of them feel economic difficult but taking economic show status, when they have the feeling difficult, everything turn out become difficult."

However, one teacher had a different view. She said,

Paragraph 2

PIT1: " This is a new subject to students, when they know it, they will like it."

ii) Students' response

When researcher asked students whether they liked economics, students that studied economics before showed their interest more than those had not study before.

During the pre-intervention interview, students stated that,

Paragraph 3

PIS8: " I like economic because I score A during secondary."
PIS5: " Yes, I like economic this subject is challenging."

On the other hand, Student 4 had contradicted opinion although she had studied economics before when she was in secondary school. She had a different belief,

Paragraph 4

PIS4: " Not really like economic. …..because that lesson is bore, if my teacher can teach it a little bit interesting, may be I will like it."

Those students that never studied economics when they were in secondary school did not showed any interest in this subject. They had a similar view with Teacher 1. A student felt that economics was extremely hard. Her interview stated as below,

Paragraph 5

PI S1: " I really doesn't know what is economic, later perhaps I will like it but not now."

C) Post- Intervention Interview

i) Teachers' perspective

It seemed the students' response of collaborative learning changed noticeably after the intervention. After the implementation of CPS, six teachers indicated students benefits from collaborative learning. They also noted that student attitude changed after the intervention.

Paragraph 6

PO*T4: " They [students] are more interested in involving in discussion. Now, they will ask for discussion during lesson. Some they have their own group outside classroom"*
PO*T1:" Most of them show more interest, except those very passive students."*

ii) Students' response

Students had a similar view with their teachers. They agreed that they could understand the economics concept easily after peer discussion. The following three comments were typically across the eight students,

POS3: "Collaborative learning really help, the problem that I could not understand, I can ask my friends. I really like economic now."
POS5: "Is really interesting to study together, we can exchange idea."
POS8: "When we try to solve the answer together, we enhance our interest."

DISCUSSION

The study found out there was a strong positive relationship between fostering students' interest on microeconomics after participation in CPS. Findings of questionnaire showed that CPS1 and CPS2 were outperformed than CG. These indicate that CPS1 and CPS2 could develop students ability of thinking which will enhance their performance. When students could understand the economics concept and score in examination, they would show interest in the subject. Lischner (2007) also agreed that collaborative method could promote students' interest. On the other hand, Evensen and Hmelo (2000) also raised the similar views. However, this findings were contradicted by Mergendoller, Lahart and Mass (2002) result. Mergendoller, et al. (2002) results failed to show any significant between interest and group discussion.

Post- interview findings revealed that students had a positive perception towards the implementation of CPS. It showed that the implementation of CPS to learn economic had encouraged students to seek answer through discussion that promote their interest. These findings also indicated that students attitude changed after the implementation of CPS. The

absence of the involvement in group learning activities might cause students lack of interest in learning economics. However, teachers have their own preferences for certain teaching method. Teachers need to consider their students interest, , ability and learning style before planning any teaching or learning method.

CONCLUSION

This study is based on pre-university students' interest after implementation of CPS. CPS is an effective learning method in fostering students' interest. At the same time, teachers must have alternative learning method in order to make their lesson lively. Besides, teachers must show their willingness to try out various learning method although they had burden with heavy workload. Lastly the positive findings supported the continuation of the future study, for example: CPS can be applied in school not only in economic but also other subjects.

REFERENCES

Becker, W., & Watts, M. (2001), "Teaching Methods in U.S. Undergraduate Economics Courses", Journal of Economic Education, 32(3), pp. 269-80.

Berg, E. (1998), "Why Aren't Aid Organizations Better Learners?" Available at: http://www.egdi.gov.sel (Accessed 4 August 2007).

Chen Yuqing, Peng Xiaoshan & Sun Jian (2010), "National Undergraduate Electronic Design Contest: A Vehicle for Enchancing Active Learning", British Journal of Educational Technology, Vol.41, No. 4, pp. 660-664.

Evensen, D.H. & Hmelo, C.E. (2000), "Problem-Based Learning~A Research Perpective on Learning Interactions", London, Lawrence Erlbaum Associates.

Goldman, R.H., Cohen, A.P. & Sheahan, F. (2008), "Using Seminar Blog to Enhance Student Participation and Learning in Public Health School Classes", American Journal of Public Health, pp.1658-1663.

Jane,D. & Jiri, D. (2009), "Use of Wiki Tools for Raising the Communicative Aspect of Learning" Proceedings of the European Conference on E-Learnig, pp.165-173.

Johnston C.G., James, R.H., Lye, J.N., & McDonald, I.M. (2000), "An Evaluation of Collaboration Problem Solving for Learning Economics", Journal of Economic Education, pp. 13-29.

Khoo, Y. Y. (2008), "The Effectiveness Collaborative Problem Solving Method Among Form Six Economic Students", Penang, University Science Malaysia.

Leung Yin Bing & Hui, A. (2009), "Creative Language Teaching: A Pilot Study Using Torrance's Incubation Model. (English)" New Horrizon in Education, Vol.57 No.1, pp.95-101.

Lischner, R. (2007), "Collaborative Learning Lite", Available at: http://www.tempest~sw.com/collab.html (Accessed 20 Jun, 2007)

Malaysia Examination Council (2008),"STPM Report", Available from: http:// www.mpm.edu.my/ (Accessed 12 April, 2008)

Mastropieri, M.A. & Scruggs, T.E. (2007), "The Inclusive Classroom: Strategies for Effective Instruction", (3rd. ed.) Upper Saddle River, NJ, Prentice Hall.

Matveev, A.V.& Milter, R.G. (2010), "An Implementation of Active Learning: Assesing The Effectiveness of The Team Infomercial Assignment", Innovations in Education and Teaching International. Vol. 47, No. 2, pp.201-213.

Mergendoller J.R., Lahart, T.E., & Mass, R.M Y. (2002), "The Effectiveness of Problem- Based Instruction: A Comparative Study of Instructional Methods and Student Characteristics" Human Investment Research and Education (HIRE) Center, California State University.

Mukhtar Ahmad Chudhary, Samina Malik, Saeed-ul-Hassan & Khalid Mahmood (2010), "An Experimental Study to Assess The Motivational Techniques Used by Teachers In The Teaching of Chemistry", Journal of Education and Sociology, pp.36-52.

Novak, J.D. (1993), "How Do We Learn Our Lesson? Taking Students Through the Process", Science Teacher, Vol.60 No. 3, pp. 50-55.

INITIAL

PIT1: Pre-Intervention Teacher 1
P1T5: Pre-Intervention Teacher 5
P1S1: Pre-Intervention Student 1
P1S4: Pre-Intervention Student 4
P1S5: Pre-Intervention Student 5
P1S8: Pre-Intervention Student 8
POT1: Post-Intervention Teacher 1
POT4: Post-Intervention Teacher 4
POS3: Post-Intervention Student 3
POS5: Post-Intervention Student 5
POS8: Post-Intervention Student 8

MOTIVATING A CLASSROOM DISCUSSION OF CENTRAL PLANNING VERSUS DECENTRALIZED MARKETS

Elizabeth A. Watson, Trinity University

John K. Watson, University of Louisiana at Lafayette

Sarah J. Skinner, University of Louisiana at Lafayette

ABSTRACT

Recent principles textbooks, in their discussion of central planning and decentralized markets, are void of details describing why central planning frequently fails. In light of recent economic crises, a strong understanding of why market organization generally leads to higher levels of GDP is essential. In order to give students this depth of understanding, we present to them data on GDP per capita, the Heritage Foundation's Index of Economic Freedom, and several indicators of resource and consumer heterogeneity. The goal of this paper is to outline the methods of presentation and classroom discussion that were used and to show, via pre- and post-test results, students gain a more in depth understanding of the material with the data presentation than with a usual "textbook" discussion alone.

KEYWORDS: *decentralized markets, central planning, comparative economic systems, teaching economics*

INTRODUCTION

It is common to find that principles of economics textbooks, in their coverage of planned versus market economies, simply define central economic planning as a system in which a central government answers the fundamental what, how, and for whom questions and then cite the collapse of the Soviet Union as evidence that central planning often fails as an economic system. The text then moves on to discuss market or decentralized economies. Very little is ever said about *why* central planning fails to produce the levels of output and growth we see in more decentralized market economies.

For instance, Frank and Bernanke (2009, 63-64) state "[w]hen implemented on a small scale, as in a self-sufficient family enterprise, centralized decision making is certainly feasible." But they offer little in the way of explaining why planning works well on small scales but not large ones. Cowen and Tabarrok (2010, 82) reason "[t]he central planning approach failed because of problems of information and incentives". Mankiw (2012, 11) expands this approach by noting "[i]n communist countries, prices were not determined in the marketplace but were dictated by central planners. These planners lacked the necessary information about consumers' tastes and producers' costs, which in a market economy is reflected in prices." Putting all these pieces together one begins to notice a more complete picture of why central planning can work on small scales whereas markets tend to be more efficient on larger ones. Information is easier to gather on small scales. Consumer preferences are more homogeneous and resources are fewer in number and scope. When there are fewer options as to what, how, and for whom to produce, the planner making optimal allocation decisions becomes much more likely. When resources and consumers are more heterogeneous, markets are necessary for efficient allocation.

However, most students will not have the incentives or the opportunities to piece together information from various texts on economics to develop the full picture. In light of the current world economic climate, it is essential to understand when and why markets work and why market organization leads to higher GDP per capita and more economic growth. For many students, even those who will go on to be business, political, and policy leaders, a principles course is as far as they will go in their formal economic education. Therefore it falls to these courses and their instructors to ensure that these future leaders understand the fundamentals of comparative economic systems. The question then becomes, how do we instill this in depth understanding of an extremely complex topic in students who have not yet taken the advanced math and statistical classes one usually associates with comparing economic growth and development across countries?

In order to spur classroom discussion on the topic and give students a more in depth understanding of the issues, the authors present students with data on GDP per capita, the Heritage Foundation's index of economic freedom (EFI), and several indicators of resource and consumer heterogeneity for various countries. The EFI gives students a measurement of the level of market decentralization in an economy while the resource variables are intended to be a proxy for the scope and scale of the economy. The goal is for students to discover, through analysis and discussion of the data, that central planning may lead to levels of GDP per capita similar to those of decentralized economies only when consumers and resources are relatively homogeneous.

In other words, central planning can work as well as decentralized markets if the task of answering the questions what, how, and for whom is simplified by having significantly less information or more homogeneous information for the planners to process. In general, a given level of planning will result in much lower GDP per capita in larger, more complex economies.

This paper presents a discussion of the data shown to students, an outline for presenting the data and guiding student discussions, and a summary of the results of in-class discussions. Using pre- and post-tests, we show students have a better understanding of the difficulties of central planning after viewing the data than they did after seeing only a "textbook" presentation of the material.

The rest of the paper is organized as follows: Section 2 gives a discussion of the data presented to students and the method used for presenting the data and guiding student discussions; Section 3 summarizes the results of in-class discussions and questionnaires; and Section 4 provides concluding remarks.

DATA AND METHODS

To begin the discussion of central planning and decentralized markets, students were presented with the textbook definitions of both types of economic organization. The students were then given a brief lecture along the lines of the Frank and Bernanke text stating central planning does tend to be efficient in very small-scale organizations but that it breaks down on larger scales due to a lack of information on the planner's part; households, small businesses, and national economies were used as examples. The students were then presented with a question, as seen in Figure 1, to test their ability to reason about the likelihood of central planning producing efficient outcomes in various countries. The data represented in the question is actual data from Qatar (Quinam) and Norway (Norland).

Given the information in the table below, for which of the two countries would central planning work better, i.e. lead to higher levels of GDP per capita? Justify your answer using the tools and terminology you learned in class. Where Education is the average years of schooling and Urban Population is the percent of the population living in urban areas.

Country	Education	Urban Population	Natural Resources
Norland	17	78	petroleum, natural gas, iron ore, copper, lead, zinc, titanium, pyrites, nickel, fish, timber, hydropower
Quinam	12	96	petroleum, natural gas, fish

Figure 1: Analysis question given to students both before and after the data presentation.

The following class period students were shown data on the Heritage Foundation's Economic Freedom Index (EFI), GDP per capita, and various indicator variables of consumer and resource heterogeneity. The EFI and GDP per capita were shown for all countries for which the Heritage Foundation calculates a Freedom Index. The other variables were shown only for those countries being used for comparison purposes. All data except the EFI was gathered from the CIA World Fact Book. At the beginning of class the graph in Figure 2 was displayed on the projector and the students were directed to note the positive relationship between a country's EFI and GDP per capita. The students were then asked if they detected any anomalies in this relationship and they quickly pointed out Luxemburg and Qatar both had higher levels of planning than several other countries but also had the highest levels of GDP per capita.

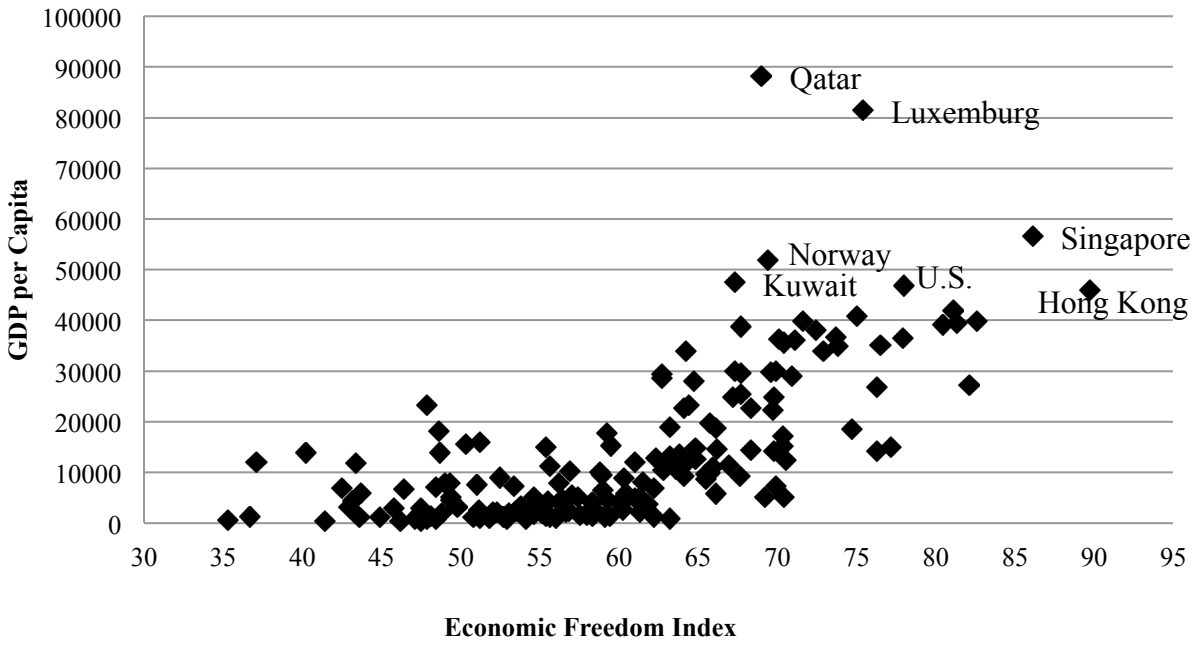


Figure 2: Scatterplot showing the relationship between GDP per capita and the Heritage Foundation's Economic Freedom Index

Once these outliers were noted, the students were shown the graphs in Figures 3 and 4. These highlight the range of the level of GDP per capita that results from roughly the same level of central planning. In particular students were directed to note the difference in GDP per capita between Qatar and Armenia and the difference between Luxemburg and Estonia. The instructor then encouraged the students to come up with explanations for why this might be the case. Initial responses were often variations of "Qatar and Luxemburg are small countries and central planning can work well on small scales". When it was pointed out that several of the countries in each subgroup with very low levels of GDP per capita were even smaller than Qatar and Luxemburg (for instance Barbados, Bahrain, and Mauritius) student explanations began to falter.

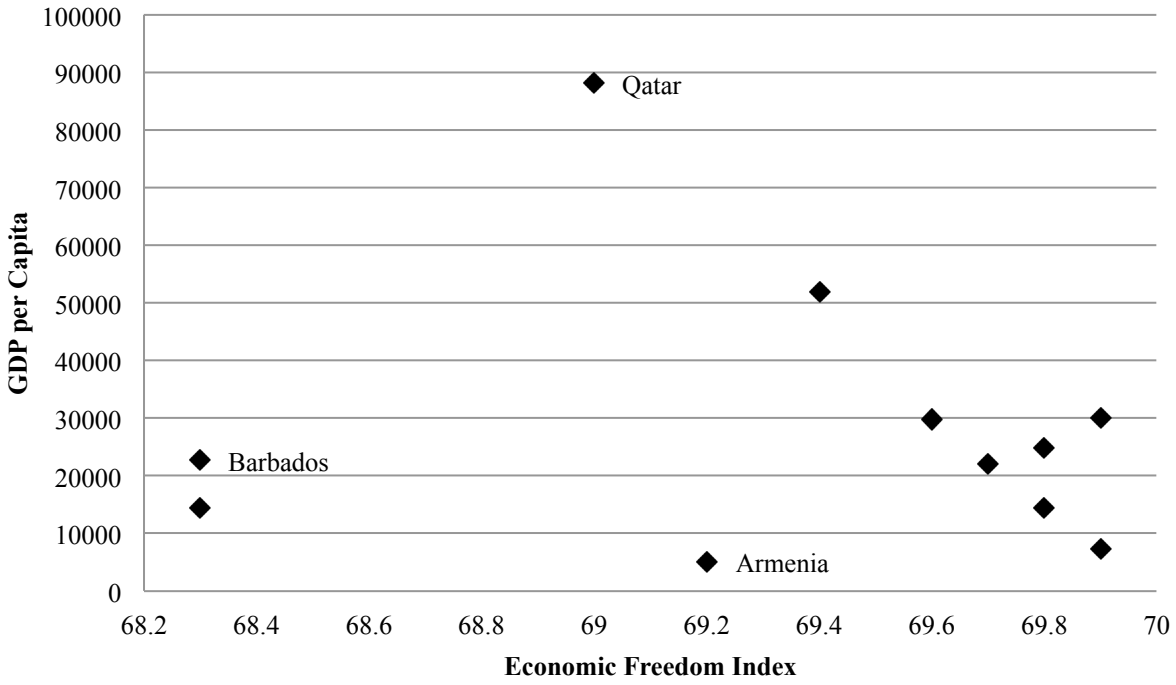


Figure 3: Scatterplot showing the GDP per capita of those countires with an EFI between 68.2 and 70.

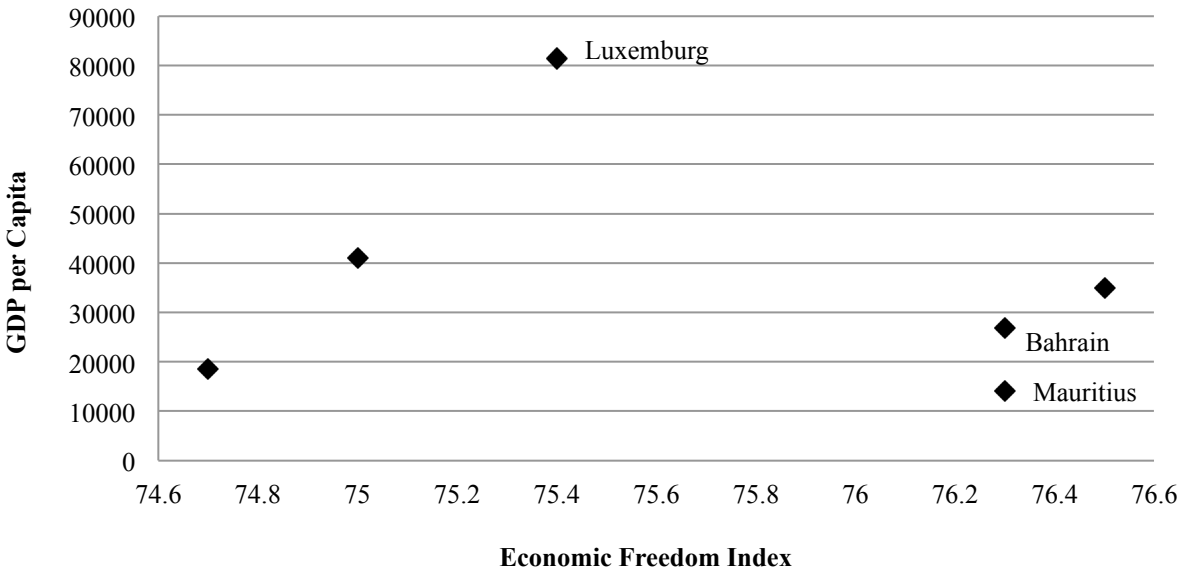


Figure 4: Scatterplot showing the GDP per capita of countries with an EFI between 74.6 and 76.6

When this happened the students were shown the data in Figure 5. This gives a breakdown of the major resources and industries (as listed in the CIA World Factbook) for each country. The students were asked to discuss why a given level of central planning would result in GDP per capita of over $85,000 for Qatar but only $5,453 for Armenia. The instructor told the students to imagine themselves as the planners and asked them to make lists of the various decisions they would have to make given the information in the table. The students quickly noted the variety of resources in Armenia and began to discuss the difficulties of trying to plan so many different types of industries. They also quickly noted the resources and industries in Qatar mostly revolved around oil. They reasoned that gathering information on a single resource would be much easier for a planner to accomplish than trying to gather information on many different resources. Once the discussion on industry planning and resource availability came to a close, the students were again given the question in Figure 1. The results of the pre-test and answers to the question were not discussed at all with the students prior to them completing the post-test. Only after all students had handed in their answers to the post-test was the question discussed in class.

Country	GDP/capita	EFI	Education	Urban Population	Natural Resources	Industries
Armenia	$5,453	69.2	12	64%	gold, copper, molybdenum, zinc, bauxite	diamond processing, metal cutting tools, forging-pressing machines, electric motors, tires, knitted wear, hosiery, shoes, silk fabrics, chemicals, trucks, instruments, microelectronics, jewelry manufacturing, software development, food processing, brandy
Qatar	$85,627	69.0	12	96%	petroleum, natural gas, fish	liquefied natural gas, crude oil production and refining, ammonia, fertilizers, petrochemicals, steel reinforcing bars, cement, commercial ship repair
Estonia	$18,410	74.7	16	69%	oil shale, peat, rare earth elements, phosphorite, clay, limestone, sand, dolomite, arable land	engineering, electronics, wood and wood products, textiles; information technology, telecommunications
Luxemburg	$82,600	75.4	13	85%	iron ore, arable land	banking and financial services, IT, telecommunications, cargo transport, food processing, chemicals, metal products (iron, aluminum, and steel), engineering, tires, glass, tourism

Figure 5: Data from the CIA World Factbook and the Heritage Foundation's Economic Freedom Index.

RESULTS

The results of the activity show a clear improvement in student understanding of the difficulties of central planning. Post-test results indicate viewing the data and engaging in class discussion about the details of the central planning process lead to more accurate reasoning about why decentralized markets lead to higher levels of GDP per capita in large-scale organizations. Table 1 presents the rubric used to score student responses to the question both before and after the data presentation. Tables 2 and 3 show a break down of student answers and rubric scores for both the pre- and post-test.

Scoring Rubric			
1	2	3	4
answer was not at all correct or simply stated that Norland would have a higher GDP because they have more resources	mentioned central planning works better on small scales but no other correct reasoning or explanation	noted that central planning can work on small scales but generally doesn't work on large scales, included some correct reasoning and explanation	noted that central planning tends to work well on small scales but not large ones, reasoning and explanation demonstrates proficient understanding of concepts

Table 1: Scoring Rubric used to analyze student answers to the pre- and post-test.

Answer	Pre-test	Post-test
Quinam	32	50
Norland	21	3

Table 2: Student answers to the first part of the question addressing for which country central planning would likely result in higher levels of GDP per capita.

Score	Pre-test	Post-test
1	23	4
2	10	8
3	11	23
4	9	18
Average	2.11	3.03

Table 3: Scores for student reasoning on both the pre- and post-test.

As shown in Table 2 there is a marked improvement in the number of students answering correctly that a given level of central planning would lead to higher levels of GDP per capita in Quinam than it would in Norland. A Chi-square test rejects the null hypothesis that the pre- and post-test answers are from the same distribution with 99.9% confidence ($X^2 = 17.5$). The data in Table 3 show a significant improvement in students' abilities to correctly reason as to why central planning would not work well in Norland, but could possibly be successful in Quinam. A Chi-square test rejects the null hypothesis that the pre- and post-test distributions are the same with 99.9% confidence ($X^2 = 20.9$).

CONCLUSIONS

Principles of economics textbooks seem to have dedicated less space to the discussion of comparative economic systems since the dissolution of the Soviet Union. Unfortunately, the time dedicated to a discussion of comparative systems in principles classes is often the only occasion that students have for formal education in this area. An understanding of when planned economies work and when they fail is essential for students who wish to grasp the nature of markets and the information that results from market activities.

Traditionally most principles textbooks have not included data in this area. Data is readily available and easily presented and the benefits to the students are clear. The findings from our classroom discussions suggest when data is presented and discussed, students develop a firmer grasp of the fact that countries that have large numbers of heterogeneous resources and consumers are more difficult to plan and therefore will have less output in the absence of decentralized markets.

REFERENCES

Baumol, William and Alan Blinder. 2010. *Economics: Principles and Policy.* 11th edition. South-Western Cengage Learning: Mason, OH

Case, Karl E., Ray C. Fair and Sharon M. Oster. 2009. *Principles of Economics.* 9th Edition. Pearson Education: Upper Saddle River, NJ

Central Intelligence Agency. 2012. *The World Factbook*. <https://www.cia.gov/library/publications/the-world-factbook/>

Cowen, Tyler and Alex Tabarrok. 2010. *Modern Principles: Economics.* Worth: New York, NY

Frank, Robert H. and Ben S. Bernanke. 2009. *Principles of Economics.* 4th edition. McGraw-Hill Irwin: New York, NY

Gwartney, James D., Richard L. Stroup, Russell S. Sobel and David A. Macpherson. 2011. *Economics: Private and Public Choice.* 13th edition. South-Western Cengage Learning: Mason, OH

The Heritage Foundation. *2011 Index of Economic Freedom*. <http://www.heritage.org/index/explore?view=by-region-country-year>

Hubbard, R. Glenn and Anthony Patrick O'Brien. 2009. *Economics.* 2nd edition. Pearson Education: Boston, MA

Krugman, Paul and Robin Wells. 2010. *Economics.* 2nd edition. Worth: New York, NY

Mankiw, N. Gregory. 2012. *Principles of Economics.* 6th edition. Thomson South-Western, Cengage Learning: Mason, OH

McEachern, William A. 2009. *Economics: A Contemporary Introduction.* 8th edition. South-Western Cengage Learning: Mason, OH

Parkin, Michael. 2010. *Economics.* 9th edition. Pearson Education: Boston, MA

Stone, Gerald W. 2011. *Core Economics.* Worth: New York, NY

CANADIAN MASTERS OF ECONOMIC PROGRAMS: HAVE THEY CHANGED OVER THE LAST DECADE

Martin Milkman, Murray State University
James McCoy, Murray State University
Riza Marjadi, Murray State University

ABSTRACT

This paper presents the results of a survey of directors of Masters of Economics programs to determine the missions, inputs, and outcomes of these degree programs throughout the United States and Canada. The survey asks respondents to provide information on program admission requirements, curriculum, faculty characteristics, enrollment, graduation rates, student financial support, and placement. The results of degree programs in the United States are compared to those in Canada. Current characteristics of Canadian degree programs are also compared to our previous survey conducted 10 years ago. Similarities and differences between programs in the United States and Canada are highlighted.

INTRODUCTION AND BRIEF REVIEW OF THE LITERATURE

This paper first presents data from terminal economics degree programs in Canada and then compares these programs to terminal economics master's degree programs in the United States. The data are from a survey of all terminal master's programs in Canada and the United States that was conducted by Survey Monkey in 2012. We then compare the results from this survey of Canadian schools to the results of a previous survey of Canadian terminal master's programs based on a survey conducted in 2002.

The first study to look at master's of economics programs was by Barr, Aby, and Willhite in 1991. They reported the results of a survey of departments offering both Ph.D.'s and master's degrees in economics. The survey focused upon the differences in the degree structures and graduate degree placement. Another study by Thornton and Inees (1988) looked at master's degrees in economics and focused on admissions requirements, curriculum requirements, and time needed to earn a degree and graduate placement into careers and doctoral programs. As noted by McCoy and Milkman (1995) "neither of these studies attempted to ascertain significant differences in programs that may be associated with different intuitional characteristics or different program goals. (p.157)" As far as we can tell the only other research in this area has been conducted in various manuscripts by McCoy and Milkman.

McCoy and Milkman published the results of an initial survey of terminal master's programs in economics in United States universities in 1995, and the results of a second survey in 2006. Both of the studies were published in the *Journal of Economic Education.* While the study published in 1995 only included programs in the United States, the survey done for the paper published in 2006 also included Canadian universities. However the results were not presented in the paper published in 2006. Instead a separate article, "Masters in Economic Programs:

Comparing Canada and the United States" was published by the *Journal of Economic Education* in 2008.

DATA AND DESCRIPTIVE STATISTICS

The data collected in 2012 was collected through an on-line survey using Survey Monkey. All of the universities in the United States and Canada listed in Peterson's Online Guide to Graduate Programs in Economics which offered a terminal master's degree in economics were surveyed. This totaled 195 in the United States and 27 in Canada. The emails were sent out in early spring of 2012. Those who did not respond were sent a second email which also contained the survey in the late spring of 2012, and this was followed by an email survey attempt to the non-responders in the fall of 2012. In the previous surveys of Masters of Economics programs the surveys were sent out by postal mail and we asked respondents to send the completed surveys back to us. The survey is very extensive and this may be why our survey response rate is slightly below what is was in the previous surveys. This time our response rate was 33 percent for the Canada universities (compared to 45 percent in the prior survey), and for United States universities the response rate was 32 percent (compared to 37 percent in the prior survey).

There seem to be less programs that offer the terminal master's degree in economics in the United States. In the United States in 2002 we were able to identify 272 programs that offered this degree, while in the 2012 survey only 195 programs were identified. However in Canada the number of programs identified only decreased by one from 34 to 33.

Table 1 and Table 2 display the results of the surveys. Table 3 contains a list and brief description of all of the variables used in the study. The survey asked respondents information about their programs such as general program characteristics, department faculty, admission requirements, student characteristics, financial aid, curriculum, and graduation rates and placement.

THE CANADIAN AND UNITED STATES PROGRAMS

Means and standard deviations are listed for all of the variables in the 2012 survey. While there are many similarities between the programs located in Canada and the United States, this portion of the paper will focus on the differences which are statistically significant.

A higher percentage of programs in Canada are more likely to see their mission as preparing students for either doctoral work or careers. Also a higher percentage of the Canadian programs are located in the Colleges of Arts and Sciences. They also have a higher percentage of women faculty members in the department.

Programs in the United States are more likely to require a standardized entrance exam. However all of the Canadian programs require a certain undergraduate grade point average for admission, while only eighty percent of the United States programs require one. The average undergraduate grade point average for the Canadian programs is 3.16 while for United States programs it is 2.982. All of the Canadian programs also require intermediate undergraduate economics.

Table 1
A Comparison of United States and Canadian Programs

	US			CA			US v. CA	
Variables[1)]	**Ma**	**SD**	**N**	**Ma**	**SD**	**N**	**z**	**Sig. [2)]**
General Program Characteristics								
PHD	0.608	0.492	74	0.8	0.422	10	-1.18	
DOCWORK	0.27	0.447	63	0.1	0.316	10	1.158	
BUSGOVT	0.635	0.485	63	0.6	0.516	10	0.212	
BOTH	0.095	0.296	63	0.3	0.483	10	-1.83	**
ARTSCI	0.466	0.502	73	0.75	0.463	8	-1.527	*
BUS	0.329	0.473	73	0.125	0.354	8	1.184	
AG	0.205	0.407	73	0.125	0.354	8	0.543	
PUBLIC	0.838	0.371	74	0.8	0.422	10	0.301	
PRIVATE	0.162	0.371	74	0.2	0.422	10	-0.301	
Departmental Faculty								
FULFAC	19.438	10.828	73	22.1	6.523	10	-1.099	
WOMFAC	18.954	13.976	62	26.292	9.457	7	-1.839	**
MINFAC	11.006	16.286	62	12.222	9.493	5	-0.258	
Admission Requirements								
ENTREXAM	0.75	0.436	64	0.1	0.316	10	4.042	***
MINIMUM	0.449	0.503	49	1	-	1	-1.094	
MINUGPA	0.794	0.408	63	1	0	10	-1.584	*
REQGPA	2.982	0.203	51	3.16	0.201	10	-2.55	***
RCRSWRK	0.875	0.333	64	1	0	10	-1.184	
PRINCECO	0.614	0.491	57	0.8	0.422	10	-1.131	
INTERECO	0.807	0.398	57	1	0	10	-1.52	*
ONECALC	0.702	0.462	57	0.7	0.483	10	0.011	
STATS	0.842	0.368	57	0.8	0.422	10	0.332	
QUANT	0.842	0.368	57	0.9	0.316	10	-0.474	
REQTOEFL	0.984	0.126	63	1	0	10	-0.401	
Student Characteristics								
FTENROLL	26.686	36.343	51	27.4	12.501	10	-0.111	
PARTROLL	11.955	45.291	44	1.571	1.618	7	1.515	*
WOMEN	35.774	20.753	44	44.4	24.228	8	-0.946	
AFRIAMER	5.734	10.058	41	1.214	2.087	8	2.604	***
HISPAMER	3.709	8.024	37	0	0	8	2.812	***
FOREIGN	38.135	26.748	42	49.374	29.044	10	-1.116	
UNDGRAD	17.748	18.037	38	32.539	59.885	10	-0.772	
UNDECON	57.452	35.114	39	78.018	31.518	10	-1.797	**
DIRECT	55.202	34.723	39	74.696	19.108	10	-2.374	***
GRADTRAN	11.412	14.495	39	4	8.966	10	2.023	**

Table 1
A Comparison of United States and Canadian Programs

	US			CA			US v. CA	
Variables[1)]	**Ma**	**SD**	**N**	**Ma**	**SD**	**N**	**z**	**Sig. [2)]**
WRKFRC	27.858	23.542	37	14.984	11.461	9	2.367	***
IN25MI	0.173	0.382	52	0	0	10	1.423	*
IN300MI	0.288	0.457	52	0.4	0.516	10	-0.701	
OUT300MI	0.538	0.503	52	0.6	0.516	10	-0.358	
Financial Aid								
TASSIST	0.561	0.501	57	1	0	10	-2.645	***
PTASSIST	19.602	29.98	49	85.5	16.236	10	-9.856	***
DSTIP-T [3)]	9812.96	4061.485	25	10179.662	5149.427	10	-0.202	
RASSIST	0.632	0.487	57	0.5	0.527	10	0.788	
PRASSIST	28.49	35.394	52	16.25	31.075	10	1.114	
DSTIP-R [3)]	9214.514	4268.247	37	8442.927	8158.127	6	0.227	
FELSCHOL	0.474	0.504	57	1	0	10	-3.087	***
DSTIP-F [3)]	6825	5534.618	11	9404.25	4970.616	8	-1.064	
OTHERAID	0.123	0.331	57	0	0	10	1.171	
NEEDBASE	0.175	0.384	57	0.1	0.316	10	0.594	
MINFAID	0.644	0.483	59	0.7	0.483	10	-0.343	
Curriculum								
RGRDHRS	32.617	7.258	47	28.125	7.68	8	1.541	*
FULLTIME	18.66	6.098	47	15	6.7	10	1.593	*
RMICRO	3.766	1.577	47	3.6	1.265	10	0.36	
RMACRO	3.283	1.515	46	3	1.414	10	0.565	
REMETRIC	4.826	3.282	46	3	2	10	2.293	**
RMATHECO	1.914	1.616	35	0	0	8	7.01	***
RSTAT	1.324	1.934	34	0	0	8	3.99	***
REQSPEC	0.277	0.452	47	0.1	0.316	10	1.178	
SPECOPT	0.419	0.499	43	0.3	0.483	10	0.691	
REQTHES	0.383	0.491	47	0.4	0.516	10	-0.1	
COMPEXAM	0.34	0.479	47	0.1	0.316	10	1.509	*
WRITEXAM	0.778	0.428	18	1	-	1	-0.531	
ORALEXAM	0.222	0.428	18	0	-	1	0.531	
Graduation Rates and Placement								
NUMBER	15.447	23.737	47	18.5	11.54	10	-0.607	
COMPLETE	83.643	12.435	42	90.9	8.319	10	-2.229	**
DOCTI	9.15	15.501	40	10.111	9.752	9	-0.236	
DOCT2	13.205	12.689	39	13.1	9.024	10	0.03	
DOCT3	3.441	3.94	34	1.429	2.44	7	1.76	**
GOVT	20.611	11.729	36	34.8	15.648	10	-2.667	***

Table 1
A Comparison of United States and Canadian Programs

	US			CA			US v. CA	
Variables[1)]	**Ma**	**SD**	**N**	**Ma**	**SD**	**N**	**z**	**Sig. [2)]**
PRIVSECT	49.649	25.044	37	35.4	17.315	10	2.08	**
POSTSEC	5.724	13.943	29	4.333	10.614	6	0.276	

Notes: 1) Variables are described in Table 3.
2) Statistical Significance * Alpha = 0.1 ** Alpha = 0.05 *** Alpha = 0.01
3) Figures are in US$

United States programs have a much higher percentage of part-time students than Canadian programs, and they have more African American and Hispanic Students. Canadian programs have significantly more students that completed an economics undergraduate major. The Canadian programs also have a higher percentage of students that entered their Masters of Economics program directly after completion of their undergraduate degree. United States programs have a higher percentage of students who transfer from another graduate program and a higher percentage of students that entered the program from the work force. Also, seventeen percent of the students in the United States programs lived within twenty-five miles prior to entering the Masters of Economics program.

All of the Canadian universities offer teaching assistantships to students in their terminal master's program, while in the United States only 56% of the programs offer this type of financial support. Over 85% of Canadian students receive teaching assistantships and in the United States less than 20% of the students in the terminal Masters of Economics programs receive teaching assistantships. All of the Canadian schools also award fellowships and scholarships to students in their programs.

The number of months it takes the typical full time student to graduate from the program is 18.66 in the United States while it is only 15 months in the Canadian programs. This may be because the Canadian programs have entering students who may be better prepared to study economics. United States programs seem to be more quantitative. The average United States program requires students to take more econometrics. None of the Canadian programs require a course in either mathematical economics or statistics. United States programs are also more likely to require a comprehensive exam. Another reason why the United States program may take longer to complete is that the average United States program has a higher credit hour requirement than the average credit hour requirement in the Canadian programs.

Almost 91% of the students in Canada complete the terminal masters of economics degree while in the United States the percentage of degree completers is 83.6%. A higher percentage of the students in Canada enter careers in government (almost 35%) after they finish their terminal Master's degree in Economics. However in the United States more students enter jobs in the private sector (almost 50%). Both the United States and Canadian programs prepare a substantial number of students for doctoral programs in economics.

Many of the differences and similarities are similar to findings in the Milkman and McCoy (2008) paper. This we believe is due to stability in the Canadian programs.

CHANGES IN THE LAST DECADE IN CANADIAN M.S. ECONOMICS PROGRAMS

Table 2 contains the means and standard deviations for the Canadian programs that responded to our surveys in 2002 and 2012. As noted above the Canadian programs have remained relatively stable, however we will highlight the statistically significant changes in the Canadian programs over the last decade.

The Canadian programs are more likely to place more emphasis on preparing students for doctoral work and preparation for careers than they did a decade ago. This may be in response to student desires since a higher percentage of the programs' graduates now take jobs in the private sector. In the 2002 survey all of the universities who responded to our survey were public institutions. In the 2012 survey 20% of the schools responding were private universities.

In the 2012 survey all of the universities responding required a minimum grade point average, while in the previous survey less than 83% had this requirement for admission. Also, while fewer universities require applicants to have a course in statistics today, they are more likely to require quantitative courses. Requiring more quantitative courses is also true of programs in the United States and this may reflect the trend of the economics profession becoming more and more mathematical.

The average number of students in each program has increased dramatically from 17.0 to 27.4 in the Canadian programs. This same trend has occurred in the United States. A greater percentage of the students in the Canadian masters programs are undergraduate economics majors. (78.01% in the 2012 survey vs. 59.86% in the 2002 survey). Fewer of the students in the Canadian programs are from public comprehensive universities.

The 2012 survey shows that a much greater percentage of students are receiving teaching assistantships than in 2002, and that the average stipend for the teaching assistantship has increased from $5,272 to $10.118. (The 2002 dollars have been adjusted for inflation.)

The percentage of students receiving research assistantships has declined over the decade, but the stipend for fellowships and scholarships has increased. The percentage of universities offering minority specific financial aid has also increased.

None of the universities in the 2012 survey required a class in mathematical economics. This may be attributed to the increasing percentages of universities that required more quantitative classes for admission. Also, fewer of the programs now require a thesis.

The percentage of students going into a doctoral program in a discipline other than economics has declined. A much greater percentage of students are now taking jobs in the private sector after they complete their master's program in economics.

In conclusion, this research has highlighted both the similarities and differences between the Canadian and United States universities terminal master's degrees in economics. It has also has illustrated the changes that Canadian universities terminal master's programs have experienced over the decade. Our hope is that information will be useful to directors of programs in both countries as they work to improve their programs.

Table 2
A Comparison of Canadian Programs

	CA 2012			CA 2002			2012 v. 2002	
Variable[1)]	**Ma**	**SD**	**N**	**Ma**	**SD**	**N**	**z**	**Sig. [2)]**
General Program Characteristics								
PHD	0.8	0.422	10	0.619	0.498	21	1.008	
DOCWORK	0.1	0.316	10	0.227	0.429	22	-0.855	
BUSGOVT	0.6	0.516	10	0.727	0.456	22	-0.72	
BOTH	0.3	0.483	10	0.045	0.213	22	2.018	**
ARTSCI	0.75	0.463	8	0.826	0.388	23	-0.469	
BUS	0.125	0.354	8	0.043	0.209	23	0.808	
AG	0.125	0.354	8	0.13	0.344	23	-0.039	
PUBLIC	0.8	0.422	10	1	0	24	-2.258	**
PRIVATE	0.2	0.422	10	0	0	24	2.258	**
Departmental faculty								
FULFAC	22.1	6.523	10	19.75	11.299	24	0.759	
WOMFAC	26.292	9.457	7	18.231	33.824	15	0.854	
MINFAC	12.222	9.493	5	24.383	33.441	15	-1.264	
Admission Requirements								
ENTREXAM	0.1	0.316	10	0.13	0.344	23	-0.246	
MINIMUM	1	N/A	1	1	0.344	23	N/A	
MINUGPA	1	0	10	0.826	0.388	23	1.407	*
REQGPA	3.16	0.201	10	3.164	0.284	17	-0.038	
RCRSWRK	1	0	10	0.957	0.209	23	0.67	
PRINCECO	0.8	0.422	10	0.682	0.477	22	0.689	
INTERECO	1	0	10	0.955	0.213	22	0.685	
ONECALC	0.7	0.483	10	0.636	0.492	22	0.351	
STATS	0.8	0.422	10	0.955	0.213	22	-1.39	*
QUANT	0.9	0.316	10	0.565	0.507	23	1.875	**
REQTOEFL	1	0	10	0.913	0.288	23	0.962	
Student Characteristics								
FTENROLL	27.4	12.501	10	17.043	12.9	23	2.166	**
PARTROLL	1.571	1.618	7	3.733	6.787	15	-1.165	
WOMEN	44.4	24.228	8	39.214	13.068	22	0.576	
AFRIAMER	1.214	2.087	8					
HISPAMER	0	0	8					
FOREIGN	49.374	29.044	10	38.504	20.215	22	1.071	
UNDGRAD	32.539	59.885	10	25.991	21.387	21	0.336	
UNDECON	78.018	31.518	10	59.86	40.643	17	1.295	*
DIRECT	74.696	19.108	10	77.472	121.54	22	-0.104	

Table 2
A Comparison of Canadian Programs

GRADTRAN	4	8.966	10	3.136	8.703	19	0.249	
WRKFRC	14.984	11.461	9					
IN25MI	0	0	10	0	0	22	N/A	
IN300MI	0.4	0.516	10	0.455	0.51	22	-0.288	
OUT300MI	0.6	0.516	10	0.545	0.51	22	0.288	
Financial Aid								
TASSIST	1	0	10	0.958	0.204	24	0.655	
PTASSIST	85.5	16.236	10	56.452	27.756	21	3.658	***
DSTIP-T [3)]	10179.662	5149.427	10	5732.99	3214.686	15	2.433	***
RASSIST	0.5	0.527	10	0.875	0.338	24	-2.349	***
PRASSIST	16.25	31.075	10	34.167	31.311	18	-1.458	*
DSTIP-R [3)]	8442.927	8158.127	6	5270.562	3160.764	21	0.933	
FELSCHOL	1	0	10	0.913	0.288	23	0.962	
DSTIP-F [3)]	9404.25	4970.616	8	5182.77	3041.426	17	2.215	**
OTHERAID	0	0	10	1	0	2	-3.464	***
NEEDBASE	0.1	0.316	10	0.261	0.449	23	-1.039	
MINFAID	0.7	0.483	10	0.13	0.344	23	3.272	***
Curriculum								
RGRDHRS	28.125	7.68	8	24.947	9.704	19	0.905	
FULLTIME	15	6.7	10	15.959	6.726	22	-0.375	
RMICRO	3.6	1.265	10	3.4	1.353	20	0.399	
RMACRO	3	1.414	10	2.789	1.032	19	0.416	
REMETRIC	3	2	10	3.5	1.762	20	-0.671	
RMATHECO	0	0	8	2.286	1.604	7	-3.771	***
RSTAT	0	0	8	0.6	1.342	5	-1	
REQSPEC	0.1	0.316	10	0.087	0.288	23	0.12	
SPECOPT	0.3	0.483	10	0.217	0.422	23	0.509	
REQTHES	0.4	0.516	10	0.913	0.288	23	-3.161	***
COMPEXAM	0.1	0.316	10	0.087	0.288	23	0.12	
WRITEXAM	1	N/A	1	1	0	2	N/A	
ORALEXAM	0	N/A	1	1	0	2	-1.732	**
Graduation Rates and Placement								
NUMBER	18.5	11.54	10	10	7.403	21	2.13	**
COMPLETE	90.9	8.319	10	84.316	10.945	19	1.81	**
DOCTI	10.111	9.752	9	17.143	22.25	14	-1.038	
DOCT2	13.1	9.024	10	18.571	17.805	14	-0.986	
DOCT3	1.429	2.44	7	5	9.405	14	-1.334	*
GOVT	34.8	15.648	10	37.143	17.944	14	-0.34	

Table 2
A Comparison of Canadian Programs

PRIVSECT	35.4	17.315	10	20.714	21.2	14	1.864	**
POSTSEC	4.333	10.614	6	1.429	3.056	14	0.659	

Notes: 1) Variables are described in Table 3.
2) Statistical Significance * Alpha = 0.1 ** Alpha = 0.05 *** Alpha = 0.01
3) 2002 figures are adjusted for inflation

Table 3
LIST OF VARIABLES

Variable	Description	Type
General Program Characteristics		
PHD	Ph.D. offered in your department?	Binary
DOCWORK	Prep for doctoral work is primary mission?	Binary
BUSGOVT	Prep for careers (bus. & gov.) is primary mission?	Binary
BOTH	Prep for doctoral work and careers are primary missions	Binary
	Administratively located in:	
ARTSCI	Arts and Sciences	Binary
BUS	Business	Binary
AG	Agriculture	Binary
PUBLIC	Public institution	Binary
PRIVATE	Private institution	Binary
Departmental Faculty		
FULFAC	Number of full-time faculty	Continuous
WOMFAC	Percentage of women faculty in department	Continuous
MINFAC	Percentage of minority faculty in department	Continuous
Admission Requirements		
ENTREXAM	Standardized entrance exam required?	Binary
MINIMUM	Required minimum score on entrance exam (if exam required)?	Binary
MINUGPA	Required minimum undergraduate GPA?	Binary
REQGPA	If required, what is the minimum GPA?	Continuous
RCRSWRK	Specific coursework required?	Binary
	If yes, is	
PRINCECO	principles of economics required?	Binary
INTERECO	intermediate economics required?	Binary
ONECALC	at least one calculus course required?	Binary
STATS	statistics required?	Binary
QUANT	Mathematical admission requirements?	Binary
REQTOEFL	Is TOEFL required?	Binary
Student Characteristics		
FTENROLL	Number of full-time students enrolled in program	Continuous

PARTROLL	Number of part-time students enrolled in program	Continuous
	Percentage of students who	
WOMEN	are women	Continuous
AFRIAMER	are African American	Continuous
HISPAMER	are Hispanic American	Continuous
FOREIGN	are foreign	Continuous
UNDGRAD	received undergraduate degrees at the same institution	Continuous
UNDECON	were undergraduate economics majors	Continuous
DIRECT	entered the program directly upon completion of bachelor's	Continuous
GRADTRAN	transferred from a graduate program	Continuous
WRKFRC	entered the program from the work force	Continuous
IN25MI	Majority of students drawn within 25 miles?	Binary
IN300MI	Majority of students drawn within 300 miles?	Binary
OUT300MI	Majority of students drawn from over 300 miles?	Binary
Financial Aid		
TASSIST	Teaching assistantships available to terminal master's students?	Binary
PTASSIST	Percentage of students in this program who receive TA's?	Continuous
DSTIP-T [2)]	Stipend of TA's	Continuous
RASSIST	Research assistantships available to terminal master's students?	Binary
PRASSIST	Percentage of students in this program who receive RA's?	Continuous
DSTIP-R [2)]	Stipend of RA's	Continuous
FELSCHOL	Fellowships or scholarships available to terminal master's students?	Binary
DSTIP-F [2)]	Stipend of Fellowships/Scholarships	Continuous
OTHERAID	Other financial aid available?	Binary
NEEDBASE	Any need-based financial aid?	Binary
MINFAID	Any minority specific financial aid?	Binary
Curriculum		
RGRDHRS	Required semester hours for master's	Continuous
FULLTIME	Number of months typically taken by full-time students to graduate	Continuous
	Number of graduate semester hours required in	
RMICRO	micro theory	Continuous
RMACRO	macro theory	Continuous
REMETRIC	econometrics	Continuous
RMATHECO	mathematical economics	Continuous
RSTAT	statistics	Continuous
REQSPEC	Is there a specialty field requirement?	Binary
SPECOPT	Is there a specialty field option?	Binary
REQTHES	Is a master's thesis required?	Binary
COMPEXAM	Comprehensive exam required?	Binary
WRITEXAM	If yes, is it written?	Binary
ORALEXAM	If yes, is it oral?	Binary
Graduation Rates and Placement		
NUMBER	How many Masters degree in this degree program are typically awarded each year?	Continuous

COMPLETE	Percentage of students who complete the program	Continuous
	Percentage of terminal master's students who upon completion begin	
DOCTI	a doctoral program in economics at same university	Continuous
DOCT2	a doctoral program in economics at another university	Continuous
DOCT3	a doctoral program in some other discipline	Continuous
GOVT	careers in government	Continuous
PRIVSECT	careers in the private sector	Continuous
POSTSEC	careers in Post-Secondary teaching	Continuous

REFERENCES

Barr, S.Z.,C.D. Aby, & S.M. Willhite (1991). Graduate economics programs: Preparation for academics or business. *Journal of Business and Economic Perspectives* 27(1), 12-17.

McCoy, J. P. & M. I. Milkman (2006). Evolution of the Masters in Economics. *Journal of Economic Education*, 37 (4), 470-476.

McCoy, J.P. & M.I. Milkman (1995). Masters in Economics Programs: Missions and Methods. *Journal of Economic Education*, 26 (2), 157-176.

Milkman, M. I. & J. P. McCoy (2008). Masters in Economic Programs: Comparing Canada and the United States. *Journal of Economic Education*, 39 (1), 92-99.

Peterson's Guide to Graduate Programs in Economics. (2008) Retrieved March 2012, from http://www.petersons.com/graduate-schools

Thornton, R.J. and J.T. Innes (1988). The status of master's programs in economics. *Journal of Economic Perspectives* 2 (1), 171-78.

PROBLEM-BASED LEARNING IN MANAGERIAL ECONOMICS WITH AN INTEGRATED CASE STUDY

Dmitriy Chulkov, Indiana University Kokomo
Dmitri Nizovtsev, Washburn University

ABSTRACT

Engaging students in managerial economics courses at the upper undergraduate and M.B.A. levels is challenging as these courses are often a theoretical standout compared with more applied business disciplines. This study describes the efforts to increase student learning with an interactive case integrated into a managerial economics course at multiple points over the semester. The underlying pedagogy utilizes principles of problem-based learning, a student-centered approach that structures learning around team-based problem solving. An analysis of assessment results suggests a positive and significant impact on student learning.

INTRODUCTION

A number of pedagogical challenges arise in teaching managerial economics courses at the upper undergraduate and M.B.A. levels. One of them is the fact that the economics component in business curriculum may be seen as a theoretical standout in comparison to the more applied business disciplines. Students may fail to see real-life application of course concepts and lose their motivation. One solution recommended in the economics education literature is case-based learning (Becker and Watts, 1995, 1998; Christensen and Hansen, 1987). A number of studies (Carlson and Schodt, 1995; Carlson and Velenchik, 2006) suggest that case-based learning (CBL) serves as a way of increasing student involvement, motivation, and learning in the economics classroom.

Another popular pedagogical strategy that originated outside of the business and economics fields and is also aimed at enhancing real-life application of theoretical concepts is problem-based learning (PBL). PBL was originally designed in medical education to address the lack of problem-solving skills in medical students (Savery 2006). In contrast to the traditional lecture-based model, PBL uses realistic problems and case studies to structure student learning around problem solving. Utilizing the PBL approach encourages students to learn not only from the instructor but also from their peers. The role of the instructor is transformed to that of a mediator. Existing empirical evidence on the impact of PBL (Dochy et al., 2003) demonstrates a positive effect on learning and problem-solving skills. Some studies in fact claim that PBL is “perhaps the most innovative instructional method conceived in the history of education” (Hung et al., 2008, p. 486).

While the implementation of CBL and such pedagogies as team-based learning (TBL) has been studied extensively in the business and economics education literature, studies on the implementation of PBL in those fields are sparse, particularly for M.B.A. courses. In this study, we present a teaching approach that integrates PBL into an M.B.A. managerial economics course at multiple points over the semester. To our knowledge, this is the first attempt to document and assess the use of the PBL approach to such a cohort. It is important to note that the M.B.A. audience is different from secondary-level or even undergraduate-level cohorts which were the focus of prior research on PBL in business and economics (Ravitz and Mergendoller, 2005; Smith and Ravitz, 2008). On one hand, M.B.A. students are usually mature, practically minded, motivated, and willing to engage, which makes them good candidates for the PBL method. On the other hand, prior studies recorded a diminishing positive effect of PBL on learning (Ravitz and Mergendoller, 2005). In other words, the biggest gain in educational goal attainment was observed among weaker students whereas the effect on stronger students was positive but insignificant. That finding may question the effectiveness of PBL in the M.B.A. classroom. The desire to examine this combination of factors has motivated our study.

The primary contribution of this study is in detailing how the PBL-based pedagogical approach is applied in the M.B.A. economics course. Furthermore, our implementation of PBL used an integrated case study to frame the students' learning experience. Our pedagogical approach thus integrated elements of the case study method with PBL. We tested this pedagogical method at two business schools over several semesters. A secondary contribution of this study is in using assessment of student learning performed in eight sections of managerial economics courses to evaluate the impact of PBL. Our analysis of assessment results suggests a robust positive impact of this integration of PBL.

The paper proceeds in the following fashion. The next section reviews the pedagogical foundations of problem-based learning. The third section describes our integration of PBL into the managerial economics curriculum. The fourth section reports the observed impact of PBL on student learning. The final section concludes with a discussion of implications for business school curricula.

PROBLEM-BASED LEARNING: A REVIEW

Problem-based learning (PBL) originated and achieved its greatest popularity in the realm of medical education. It was developed in the latter half of the 20th century based on ideas advocated by various education and psychology researchers (e.g. Bruner, 1959, 1961; Dewey, 1910; Rogers, 1969). PBL served as a response to criticisms of traditional lecture-based education in the medical field (Barrows, 1996). In recent years, PBL has seen applications in a variety of secondary and higher education settings (Hung et al., 2008). The essence of the PBL approach is to let students learn by solving realistic problems that may not have a uniquely correct answer. This provides the students an opportunity to solidify their knowledge of theoretical concepts by trying out various solutions without being punished for mistakes (Gijselaers et al., 1995).

A basic description of the PBL approach presented by Barrows (1996) as well as Hung et al. (2008) involves the following elements. First, unstructured questions or problems are assigned to groups of five or more students. They work to define and bound the problem based on what they already know, and develop hypotheses or conjectures that lead them to identify what they need to find a solution. Next is the self-directed study stage in which individual students or the entire group complete their learning assignments. Individual results are reported back to the group. The instructor or tutor serves as a facilitator who supports reasoning and helps organize group and interpersonal dynamics, rather than provides direct answers to student questions. Finally, at the end of the learning period students summarize and integrate their findings and solutions.

Overall, PBL is a structured pedagogical approach that is distinct from simple problem-solving embedded in a traditional course. When the solution mechanism is well-defined as in traditional textbook problems, there is little to no pressure on the learner to formulate their own approach to finding a solution. Note that in PBL this role of the student is enhanced as problems are not clearly defined at the outset. The student-centered approach in PBL requires that the students find the information and techniques necessary for the solution on their own. As a result, problem-solving in PBL often replaces traditional lectures and serves as the main mechanism of learning (Savery, 2006).

PBL is also distinct from other popular group pedagogies such as case-based learning (CBL) and team-based learning (TBL). CBL combines two elements: the case itself and the classroom discussion. A teaching case is a rich narrative in which individuals or groups must make a decision or solve a problem. This narrative provides information, but not analysis. The analysis of the events in the case, identifying options, evaluating choices and consequences of actions is performed by students during the classroom discussion (Carlson and Schodt 1995; Carlson and Velenchik 2006).

There are a number of similarities between CBL and PBL as both methods include student-centered learning facilitated by the instructor. The differences lie in the focus of these pedagogies. The case method focuses on real-life examples, whereas PBL problems, while being realistic, are often synthesized. For CBL implementations the group size is of secondary importance, and the discussion - a key element of CBL pedagogy - typically involves the entire class. Meanwhile, PBL implementations typically involve groups of around five students, and group-level discussion is often more prominent than the classroom-level one since the group assumes responsibility for their findings.

There are also similarities and differences between PBL and team-based learning (TBL). TBL is another teaching method that has been used effectively both in medical and science education (Michaelsen et al., 2002) as well as business schools (Hernandez, 2002). In TBL, students cover the assigned reading prior to class and then work in groups to solve problems by applying the knowledge they acquired in the readings (Fink, 2002; Parmelee and Michaelsen, 2010). These problems may often incorporate case studies that allow the students to see a realistic application of their analysis (Quarstein and

Peterson, 2001). Peer pressure is a characteristic of TBL, as students are held accountable by their group members for absences and failure to perform (Kreie et al., 2007). While both TBL and PBL use group dynamics and peer feedback to promote learning, there is a distinct difference as the group work in TBL involves traditional structured case studies and problems covered after the assigned readings or lectures. In contrast, PBL requires the groups to determine what information or content is necessary to learn in the process of problem-solving and presents unstructured problems.

While PBL is an established pedagogy in medical education, studies on the effectiveness of PBL implementation have become increasingly popular outside of the medical disciplines (Barrows, 2000; Gallagher, 1997; Hmelo-Silver, 2004; Tan and Hung, 2007). Dochy et al. (2003) conclude based on a meta-analysis of 43 empirical studies of PBL that there is a robust positive impact on students' ability to apply their skills and the retention of learning. Dochy et al. (2003) report that students in PBL classrooms gained slightly less knowledge but retained their knowledge better. Similarly, Norman and Schmidt (1992) conclude that PBL might not improve the initial acquisition of knowledge by the students while the deeper processing of information in PBL classes leads to better retention of knowledge over a longer period of time.

In the business and economics education, the implementation of PBL has not been as common. While Gijselaers et al. (1995) as well as Stinson and Milter (1996) propose PBL as a way to enhance problem-solving skills in business students, fewer empirical studies on the impact of PBL in business education exist. Maxwell et al. (2001) report that PBL may be an effective way to engage students in high-school microeconomics classes. Meanwhile the studies of Mergendoller et al. (2000), Ravitz and Mergendoller (2005), and Smith and Ravitz (2008) report mixed evidence on PBL effectiveness in comparison to traditional lecture- and discussion-based economics classes.

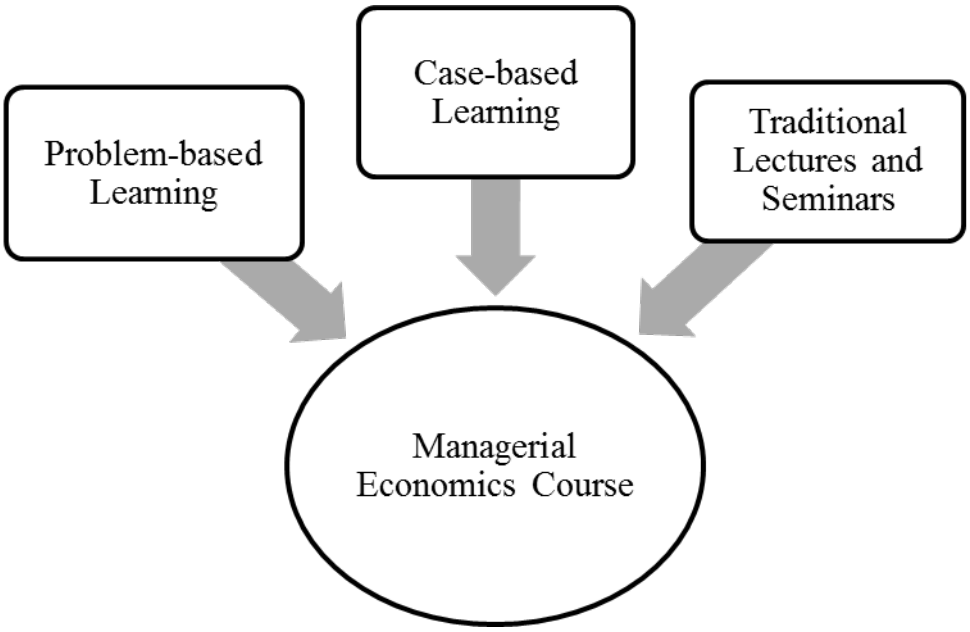
INTEGRATION OF PROBLEM-BASED LEARNING INTO MANAGERIAL ECONOMICS CURRICULUM

The managerial economics component in a typical undergraduate or graduate business program is often presented in a stand-alone course. The characteristics of the M.B.A. student population along with the more theoretical focus of such a course relative to other business disciplines made us consider utilizing PBL principles in the M.B.A. economics classroom. At the same time, the need to address a list of pre-set learning outcomes in a managerial economics course places a constraint on the implementation of pure PBL in such a course. As noted by Norman and Schmidt (1992) as well as Dochy et al. (2003), PBL classes typically are not designed to take students over a long list of content-related learning outcomes. We addressed this challenge by adopting a pedagogical strategy that supplemented PBL with an integrated case study and traditional lectures. Our pedagogical approach is illustrated by Figure 1.

In order to improve student experience with PBL, we decided to structure the PBL component of the course around an integrated case study that focuses around a common theme and presents a series of ten PBL assignments. We used the modular case

assignments published in Chulkov and Nizovtsev (2012) and organized PBL coursework around these. Each assignment in this case looks at a problem faced by the management of a car rental agency serving a small regional airport in a US Midwestern college town. The assignments are unstructured problems presented in the form of 'memos' simulating communication with the upper management.

Figure 1



The assignments in the case are both quantitative and qualitative. Examples of qualitative questions include such issues as advertising media choices or critical analysis of price match guarantees and price discrimination. Quantitative assignments ask students to perform statistical analysis for demand estimation, and use those results for forecasting and optimization. This case is well-suited for PBL because the student groups need to organize the analytical process on their own and perform such tasks as selecting appropriate variables for analysis and analytical techniques. The instructor provides only a subset of the entire case dataset to each group. In the end, no two groups have exactly the same sample of data and exactly the same results of the analysis. This provides the instructor with the opportunity to teach the methodology of analysis in managerial economics by comparing various approaches and their effect on results.

In order to integrate PBL into this managerial economics course, we reviewed the list of the eight learning outcomes for the course and identified ones to be used for our implementation of PBL. While it is possible to create PBL assignments for all the learning outcomes in a managerial economics course, this implementation covered only five of the learning outcomes presented in Table 1.

The three learning outcomes not included in the PBL component of the course were covered exclusively with traditional lectures and discussion. Meanwhile the other

five learning outcomes, such as outcome 1 “Demonstrate knowledge of optimization techniques” and outcome 2 “Apply regression tools to the analysis of real business problems” were the focus of the PBL. Following brief traditional lectures on these topics, the students were presented with the case assignments. They were organized in small groups. As required by the PBL process, the groups needed to structure and bound the problems presented to them, determine the information needed to work out a solution, select analytical methods such as optimization or regression analysis, and then organize the process of solving the problem. Each case assignment was discussed in more than one class period, so the groups often divided the work and then brought together the solutions developed by different group members. The students also searched the textbook and additional learning resources as they were working on their solutions. Each group worked out a final solution to the case assignments and presented these in the appropriate format for a business audience. The final solutions were presented to the other groups in the course, which provided the instructor and the students with the opportunity to focus on the analytical processes used and highlight the role of various assumptions and analytical techniques.

Table 1
LEARNING OUTCOMES IN THE MANAGERIAL ECONOMICS COURSE

	Learning Outcome	Included in PBL Implementation
1	Demonstrate knowledge of optimization techniques	Yes
2	Apply regression tools to the analysis of real business problems	Yes
3	Explain variations in market structure across industries and the effect of market structure on optimal behavior of firms	No
4	Analyze pricing strategies used by firms in the marketplace	Yes
5	Model strategic interactions between firms in the marketplace using tools of game theory	No
6	Apply principles of making business decisions under uncertainty	No
7	Analyze diverse and unstructured real-world problems and cases using cost-and-benefit analysis and marginal analysis	Yes
8	Critically and objectively evaluate decisions made by businesses and policymakers	Yes

This pedagogical method based on PBL structured around an integrated case study has been applied in M.B.A. managerial economics classes at two business schools over several semesters with total class sizes ranging from fifteen to thirty-six students. The size of the teams working on PBL assignments ranged from three to five students. An important component of PBL is peer feedback, which was provided during presentations of teams’ findings at several points in the semester. As the teams worked on the assignments, the instructor also provided feedback on each submitted or presented

part of the overall case. The final grade for the PBL project was based on the final written summary of all answers to the case assignments due at the end of the semester.

IMPACT ON STUDENT LEARNING

The impact of PBL on student learning was evaluated as part of assessment of student learning, in which the attainment of course learning outcomes was tested on a comprehensive final exam given at the end of the semester. The advantage of this approach was that the course learning outcomes, and the mapping of these outcomes to specific assessment instruments could not be influenced by the researchers. Both were created for the general assessment process within the assurance of learning system that was present in the two accredited business schools that were part of this study. We use these existing assessment data to evaluate whether cohorts that were subject to PBL exhibit any systematic differences in the attainment of course learning outcomes. A limitation of the study is posed by the fact that the assignment of students to PBL and non-PBL treatments was based on their course section cohort, making it a quasi-experiment rather than a pure randomized trial.

Each course learning outcome was assessed using a mix of quantitative, short-answer, and multiple-choice questions appropriate for that outcome. Student performance was measured as the percentage of the maximum possible score for each learning outcome. Assessment data were collected from a total sample of 160 students representing eight different sections of M.B.A.-level managerial economics courses taught between 2008 and 2012. Four of these sections with a total of 65 students relied on the traditional lecture-based approach, while the other four sections with a total of 95 students utilized PBL throughout the semester. Both institutions were equally represented in the sample. Throughout the period covered by the study, both the completely lecture-based sections and the PBL-based sections of the courses were taught by the same two instructors in the same proportion. The mix of assessment questions and the grading rubrics did not vary across institutions or course sections. Test papers from different sections were mixed, and each test was graded twice independently by two different graders.

In order to determine whether the introduction of PBL affected the attainment of learning outcomes by students, the two-tailed heteroscedastic t-test for equality of sample means (Welch's t-test) was performed on samples of individual student assessment scores recorded in the sections with and without PBL implementation. As mentioned above, these assessment scores were reported as a percentage of the maximum possible score on the section of the exam linked to a specific learning outcome. There were five learning outcomes that were featured in the PBL assignments in the course sections that utilized PBL, and three learning outcomes that were only featured in traditional lectures in all of the sections. The null hypothesis was that the two sample mean scores were equal, and the alternative hypothesis was that the mean score for the sample that used PBL was different from the sample that did not use PBL. The descriptive statistics and results of this analysis are shown in Table 2 and Figure 2 below.

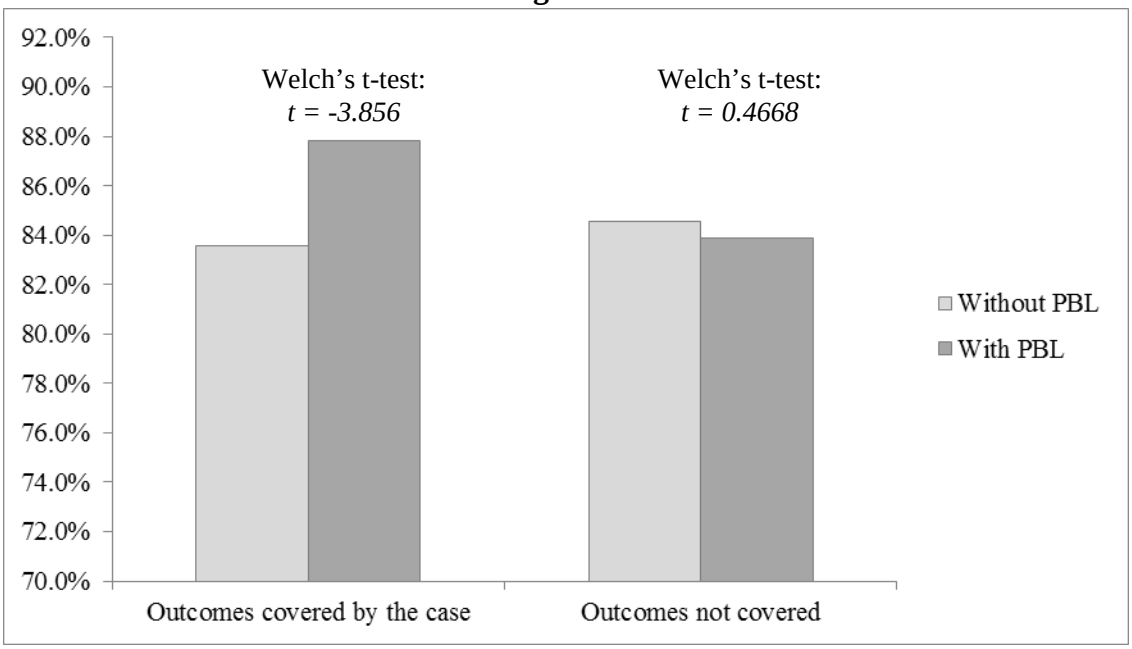
Table 2
IMPACT OF PBL ON STUDENT PERFORMANCE

	Performance scores for five learning outcomes covered by PBL		Performance scores for three learning outcomes not covered by PBL	
	Section with no PBL	Section with PBL	Section with no PBL	Section with PBL
Mean score	0.836	0.879	0.845	0.839
Variance	0.023	0.023	0.020	0.031
Observations	325	475	195	267
Degrees of freedom	687		457	
t-statistic	- 3.856**		0.467	
p-value	0.0001		0.6409	

*p< .05; **p< .01 Two-tailed heteroscedastic t-test was used.

Our findings indicate a significant variation in the way PBL affected student learning. The mean performance for the three learning outcomes not addressed by the PBL approach in either of the course sections did not change significantly between sections. Based on the result of the Welch's t-test, we are unable to reject the null hypothesis about the equality of the two sample means for these data. For those three learning outcomes, the presence of PBL in the course did not improve average student performance.

Figure 2



The results for the five learning outcomes addressed by the PBL component of the course are, however, strikingly different. For these five outcomes, the null hypothesis about the equality of sample means from sections that used PBL and sections that did not

use PBL is confidently rejected. Average student performance metrics are significantly higher for sections in which PBL was used as seen in Table 2 as well as Figure 2 below.

Overall, these findings suggest that improvement in the attainment of student learning outcomes correlates with the utilization of the PBL-based pedagogical method. Student subjective satisfaction with the course as reported by end-of-semester evaluations has also been higher in the sections that utilized PBL.

CONCLUSION

Problem-based learning (PBL) uses realistic problems and case studies to embed student learning into problem solving. PBL pedagogy encourages students to learn not only from the instructor but from their peers as well. Learning occurs as students bound and define an unstructured problem, select proper problem-solving techniques, and present solutions. These characteristics make PBL a potentially powerful tool for improving student engagement and content retention in managerial economics courses, especially at the M.B.A. level.

While PBL is a well-known pedagogical method in a number of academic areas (Hung et al., 2008), its impact on business and economics education particularly at the M.B.A. level is not well documented. This study attempts to fill this gap. We describe an implementation of PBL within the M.B.A. managerial economics curriculum at two business schools. Our implementation of PBL was centered around an integrated case study that presented ten PBL assignments at various points of the semester. Chulkov and Nizovtsev (2012) present the full text of the case study assignments used in this integration of PBL. Analysis of assessment of student learning data demonstrates a robust positive impact of this integration of PBL in the managerial economics curriculum.

Overall, we find the results encouraging. They suggest it is possible to replace some of the traditional assignments and delivery methods with cases optimized for PBL and by doing so improve the student knowledge and skills achieved in the course. Instructors teaching managerial economics courses are encouraged to consider this approach for integrating PBL into the classroom and engage in an active discussion of PBL ideas in order to plan and enact PBL assignments effectively.

REFERENCES

Barrows, H. (1996). Problem-based learning in medicine and beyond: a brief overview. In L. Wilkerson & W. Gijselaers (Eds.) *Bringing problem-based learning to higher education: Theory and practice*. San Francisco, CA: Jossey-Bass.

Barrows, H. (2000). *Problem-based learning applied to medical education*. Springfield, IL: Southern Illinois University.

Becker, W., Watts, M. (1995). Teaching tools: Teaching methods in undergraduate economics. *Economic Inquiry, 33*, 692–700.

Becker, W., Watts, M. (1998). *Teaching economics to undergraduates: Alternatives to chalk and talk.* Cheltenham, UK and Northampton, MA: Edward Elgar.

Bruner, J. (1959). Learning and thinking. *Harvard Educational Review, 29*, 184–192.

Bruner, J. (1961). The act of discovery. *Harvard Educational Review, 31*, 21–32.

Carlson, J., Schodt, A. (1995). Beyond the lecture: Case teaching and the learning of economic theory.

Journal of Economic Education, 26, 17-28.
Carlson, J., Velenchik, A. (2006). Using the case method in the economics classroom. In W. Becker, M. Watts, S. Becker (Eds.). *Teaching economics: More alternatives to chalk and talk*. Cheltenham, UK and Northampton, MA: Edward Elgar.
Christensen, C., Hansen, A. 1987. *Teaching and the case method*. Cambridge, MA: Harvard.
Chulkov, D., Nizovtsev, D. (2012). Rent-A-Car: an integrated team-based case study for managerial economics. *Journal of Business Cases and Applications, 6*, 1-14.
Dewey, J. (1910). *How we think*. Boston: Health & Co.
Dochy, F., Segers, M., Van den Bossche, P., Gijbels, D. (2003). Effects of problem-based learning: a meta-analysis. *Learning and Instruction, 13*, 533–568.
Fink, L. (2002). Beyond small groups: Harnessing the extraordinary power of learning teams. In L. Michaelsen, A. Knight, L. Fink (Eds.) *Team based learning: a transformative use of small groups in college teaching*. Sterling, VA: Stylus Publishing.
Gallagher, S.A. (1997). Problem-based learning: where did it come from, what does it do, and where is it going? *Journal of Education for Gifted, 20*(4), 332–362.
Gijselaers, W., Tempelaar, D., Keizer, P., Blommaert, J., Bernard, E., Kasper, H., (1995). *Educational innovation in economics and business administration: The case of problem-based learning*. Norwell, MA: Kluwer.
Hernandez, S. (2002). Team learning in a marketing principles course: Cooperative structures that facilitate active learning and higher level thinking. *Journal of Marketing Education, 24*(1), 73-85.
Hmelo-Silver, C. (2004). Problem-based learning: what and how do students learn? *Educational Psychology Review, 16*(3), 235–266.
Hung, W., Jonassen, D., Liu, R. (2008). Problem-based learning. In J. Spector, J. van Merrienboer, M. Driscoll (Eds.) *Handbook of research on educational communications and technology*, 3rd Ed. New York: Lawrence Erlbaum Associates.
Kreie, J., Headrick, R., Steiner, R. (2007). Using team learning to improve student retention. *College Teaching*, *55*(2), 51-56.
Maxwell, N. L., Bellisimo, Y., and Mergendoller, J. (2001). Problem-based learning: modifying the medical school model for teaching high school economics. *Social Studies, 92*(2), 73–78.
Mergendoller, J. , Maxwell, N. , Bellisimo, Y. (2000). Comparing problem-based learning and traditional instruction in high school economics. *Journal of Education Research, 93*(6), 374–382.
Michaelsen, L., Knight, A. and Fink, L. (2002). *Team based learning: a transformative use of small groups in college teaching*. Sterling, VA: Stylus Publishing.
Norman, G., Schmidt, H. (1992). The psychological basis of problem-based learning: a review of the evidence. *Academic Medicine, 67*(9), 557–565.
Parmelee, D., Michaelsen, L. (2010). Twelve tips for doing effective team-based learning. *Medical Teacher, 32*(2), 118-122.
Quarstein, V., Peterson, P. (2001).Assessment of cooperative learning: a goal-criterion approach. *Innovative Higher Education*, *26*(1), 59-77.
Ravitz, J., Mergendoller, J. (2005). Evaluating Implementation and Impacts of Problem Based Economics in U.S. High Schools. *Annual Meeting of the American Educational Research Association*, Montreal, Canada.
Rogers, C. (1969). *Freedom to learn*. Columbus, OH: Merrill Publishing.
Savery, J. (2006). Overview of problem-based learning: Definitions and distinctions. *Interdisciplinary Journal of Problem-based Learning, 1*(1), 9-20.
Smith, T. Ravitz, J. (2008). Problem Based Learning in College Economics. *Academic Exchange Quarterly, 12*(1): 22-28.
Stinson, J., Milter, R. (1996). Problem-based learning in business education: Curriculum design and implementation issues. In L. Wilkerson & W. Gijselaers (Eds.) *Bringing problem-based learning to higher education: Theory and practice*. San Francisco, CA: Jossey-Bass.
Tan, O., Hung, D. (2007). *Problem-based learning in e-learning breakthroughs*. Singapore: Thomson.

EDUCATION AND ITS CONTRIBUTIONS TO THE ECONOMY

Ashley Ardiente, University of Guam
Jonah May V. Guiking, University of Guam

ABSTRACT

Education demonstrates its importance to economic growth as a contributor to human development, production quality, and technology advancement. Moreover, this research aims to explain and compare the quality of education between least developed countries (LDC) such as South Africa and newly industrialized countries (NIC) such as Taiwan. This study will explain the factors that contribute to the lack of proper education and other hardships in South Africa that prevent its people from acquiring even the minimum of primary schooling. Then we will take a look at the NIC, Taiwan and analyze its success with education. A comparison will be made with Taiwan's experience with its education reform and the effects it had on their economy to South Africa's situation. With this comparison, we hope to provide suggestions and ideas for LDCs to consider improving their education system which in turn will help their country's economy grow.

INTRODUCTION

Having a good education can be a very valuable asset that can improve and enhance the quality of one's life. According to Glewwe et al. (2011), "economists and other researchers have accumulated a large amount of evidence that education increases workers' productivity and thus increases their incomes" (p. 1). Education can help in reaching success and higher goals in life. By having education, one has the greater advantage of knowledge that can help them surpass the common rest. Obtaining a higher education can help one get a higher level job and help reduce the unemployment rate in their country. Furthermore, Hanushek and Woessmann (2008) state that there is also a huge amount of indications that prove that education has a positive effect on economic growth (Glewwe et al., 2011). When people are educated, they provide better services to their country which then can attribute to the economic growth of their land. Education does not only benefit the individual, but also everyone as a whole.

Based on all these information and confirmed results, there is a great importance in education. Countries can examine this to help improve their economic state. This paper will look at two different regions, South Africa and Taiwan to give a comparison on the educational state of their countries. By looking at these two regions, the educational system of a least developed country (South Africa) will be compared to a newly industrialized country (Taiwan) to determine the true effects that education has on the economy.

SOUTH AFRICA (LDC)

With this being said, education has proven to be one of the influencing areas to the positive contribution to a country's economy. With this fact, countries should put a focus and goal to improve their education system because although it can be very costly, it is an investment

that will yield a better return in the future. The Reviews for National Policies for Education: South Africa (2008) writes that "it makes schooling compulsory for all children from the year they turn 7 to the year in which they turn 15 (or the end of grade 9, whichever comes first)" (p. 39). However, even with this implemented compulsory education, there are still millions of kids who do not attend school. Kennedy (2007) states that "some 46 million African children — nearly half the school-age population — have never set foot in a classroom, according to the United Nations" (para. 5). But why is this so when children are mandated to attend school?

The governments of South Africa have continually tried to improve and develop a better education system. However, although South Africa has already and is still continuing their efforts in providing education for their people, there are still many obstacles that its people face that prevents them from attaining this goal. The government of South Africa according to the Education Challenges in South Africa and LDCs (2011) writes that African governments have made continuous efforts such as developing the Sector Wide Approach (SWAp) program, improving literacy in South Africa, and abolition of primary school fees throughout the country. Also a report by UNESCO states that over the past ten years, the budget for education has increased by six percent annually. However, even with all these programs less developed countries like South Africa face more difficult complications and challenges as they try to improve their country's education that can hopefully one day benefit and help their country's economy. Some of the hardships Africans face with sending their children to get a higher and better education is the costs, quality of education they receive, impacts of family attributes to the pursuit of education, and the difficulty of retaining the children to continue their education.

One of African's biggest challenge and continuing issue is retaining the school enrollment. Based on the recent Education Statistics in South Africa 2012 (2014), "the average test score for Grade 9 learners was 43.4% in Home Language, 34.6% in First Additional Language and 12.7% in Mathematics. Only about 2% of learners obtained 50% and more in Mathematics" (p. 45). Below is also a table that summarizes the educational status of South Africa:

Table 1: Average Test Scores

	Average Percentage Mark	Percentage of Learners Achieving 50% or More	Average Percentage Mark	Percentage of Learners Achieving 50% or More	Average Percentage Mark	Percentage of Learners Achieving 50% or More
	Home Language		Mathematics		First Language	
Grade 1	57.5	63.7	68.1	77.4	-	-
Grade 2	55.3	64	57.4	67.8	-	-
Grade 3	52	56.6	41.2	36.3	-	-
Grade 4	42.6	41.3	37	26.3	33.6	25.2
Grade 5	39.9	36.5	30.4	16.1	29.6	15.9
Grade 6	42.8	38.7	26.7	10.6	35.6	24.4
Grade 9	43.4	38.9	12.7	2.3	34.6	20.8

Source: Education Statistics in South Africa 2012 (January 2014) by the Department of Basic Education.

This shows that even those who continue their education past primary school do not have adequate and efficient knowledge. It also shows that as children progress through their education, their ability to retain the knowledge actually decreases instead of the expected increase.

The first and most influential factor that prevents most children from getting an education in South Africa is the monetary costs of going to school. Although education is offered for free, like the public school system, there are other hidden costs that must be paid in order for the child to have an adequate education. According to Epstein and Yuthas (2012), education might be free but expenses for lunch, uniforms, school fees, travel and transportation costs, etc. must be paid out of the pockets of the individuals. Moreover, because quality of education in these areas are low, parents also pay tutoring fees for their children to pass examinations. Furthermore, Kennedy (2007) writes that "60,000 students who had been admitted to public secondary schools this year failed to report because they would still be charged for transport, field trips, teacher conferences and building swimming pools for schools" (para. 15). Kennedy (2007) also continues by saying that fewer than half continue on to secondary school because fees can result up to $100 while the average annual income of families is $300. Besides the extreme costs of education, there is also an opportunity cost lost—the children could have acquired jobs to help support their families instead of going to school to get inadequate education that does not even provide them basic literacy or numeracy. Just with these factors alone, parents will be more discouraged to put a child to school when it can bring more complications and extra costs that add up to the things they already cannot afford.

Another major factor that LDCs face in promoting education is providing quality learning and environment. To provide the adequate education, instructors should also be qualified and have the proper skills and abilities. According to many research and other case studies that have been published, quality of teacher education has yielded different results. Meier and Rauch (2005) state that out of sixty-three studies performed that considered the input of teacher's education, thirty-five studies were proven to have a positive impact to the quality of education they provide while only two studies showed a negative effect. However, even with this high difference, there were twenty-six studies that state that the teacher's education had a statistically insignificant impact on the quality of education.

Many schools might be built in different communities, but if there is a lack of supplies and materials or if the teacher themselves are inadequate, then how can the country provide quality education that is of value and worth and that will benefit the child? According to the Education Challenges in South Africa and LDCS (2011), it writes that "in 2007, the pupil-to-teacher ratio at the primary level increased by nearly 2 percent to 46.6 percent. With this rate, additional 1.2 million qualified teachers are required to ensure a good learning environment for all" (p. 2). Additionally, the Education Statistics of South Africa 2012 (2014) report also show the current ratios of education in the table below:

Table 2: Education Statistical Ratios

South Africa (2012)	
Student-Teacher Ratio	29.2:1
Student-School Ratio	481.0:1
Teacher-School Ratio	16.5:1

This information shows that even if children are attending school and pursuing an education, there is a great lack of supply of teachers to feed the demand. Because of the inadequate amount of teachers, there is also an insufficient amount of education being provided.

With all these information, data, and statistics, it is very evident why LDC's, especially South Africa, have a very difficult road in trying to fix their educational system. But South Africa should not lose hope. There are several other LDC's that were in a similar state, but with

strategic endeavors and preservation, they were able to surpass this unfavorable situation. One very good example of such is Taiwan.

TAIWAN (NIC)

Taiwan went through an economic transformation. In the 1950's, Taiwan still fell under the category of being a less developed country that experienced poor economic conditions, overpopulation, and scarce resources. In the 1960's Taiwan started to improve its agriculture which led the country to become more industrial. In addition to Taiwan's changes, economists believed that most of the country's economic growth stemmed from the improvement of education. Over four decades the economic growth was 8.45% between 1960 and 2000 (Chaung & Lai). The average years of education completed by the employees increased over the years. Starting from 1960's, the average years of education completed was 3.57 then by the 2000's it increased to 10.86 years (Godo, 2012). In addition, during the period of 1978 to 2006 the per capita income increased from USD 1,461 to USD 14,455 (Chaung & Lai). Human capital accumulated more over the years through the implementation of the Taiwan's government policies such as the education reformation in 1968 (Lin, 2003).

After World War II, the country's agriculture suffered a decline due to war damages and terrible weather condition. The main focus in the development of Taiwan was to improve its agriculture and regain economic stability. In order to do so, the government provided incentives for farmers to improve agriculture production (Mao & Schive). The improvement on agriculture developed a path towards industrialization. Taiwan's plan was to build the industry through agriculture and develop their agriculture through industry. This meant that they wanted to build a balance between the two to improve their economy (Mellor, 1995). The expansion of Taiwan's agriculture provided products for their country as well as the opportunity to export their surplus to other countries in order to benefit from foreign trade. As Taiwan began to develop its industry, the government turned their focus onto education. Taiwan needed to be able to provide adequate employees for their labor-intensive society. Therefore, Taiwan then shifted its focus towards the improvement of human capital. The economic return on education would not only impact the individual, but also the productivity in the society.

The government pressed forward for the improvement on the education and established the nine-year compulsory education program in 1968. The changes in society made it essential for the government to focus on providing education to learning basic and advanced skills. Taiwan wanted to focus on the problems that the younger generations will be facing such as technology advancement, so the future of the society would improve (Li, 1999). Taiwan's Ministry of Education wanted to ensure the quality of primary education for the students, so they provided the option of enrolling children into pre-school and kindergarten to prepare students before they enter first grade. The Ministry of Education also limited the class sizes for an adequate learning environment, provided English learning classes, offered internet access, and enhanced teacher training programs (Li, 1999). Taiwan was not only trying to improve the education of students through the offered programs, but through enhancing the talent of the teachers. Another implementation that the Ministry of Education applied was the assigning of students to classes accordingly. Schools are allowed to place the students according to interests, attitudes, and skills of the students. By doing so, they prepare the students to enroll in vocational or general high schools to develop suitable skills for their future career. The encouragement in continuing to pursue a higher education deemed successful in Taiwan's society. Taiwan entered

into the new era of globalization and became a knowledge-based economy. In 2012, Taiwan had 163 universities. Higher education became more available that is was not only for the elite few, but for the general public who wanted to continue their education (Higher Education in Taiwan, 2012).

The following two tables (Tables 3 and 4) consist of data collected from 1960 to 2000 of Taiwan's population. Table 3 shows the average increase in the number of years of schooling. Table 4 shows the growth of real gross domestic product in intervals of ten years.

Table 3: Increase in Years of Completed Education

Year	Average Years of Schooling
1960	3.57
1970	5.09
1980	7.32
1990	9.12
2000	10.86

Source: Y. Godo, A New database on Education Stock in Taiwan (2012)

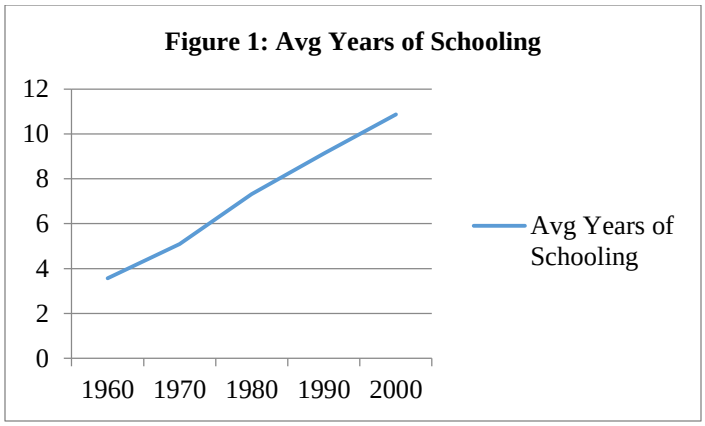
In Table 3, the average number of years of attending school was significantly low in 1960. However, once the education reform began in 1968 with implementing compulsory education, people attended school longer which meant that more of the population was being educated.

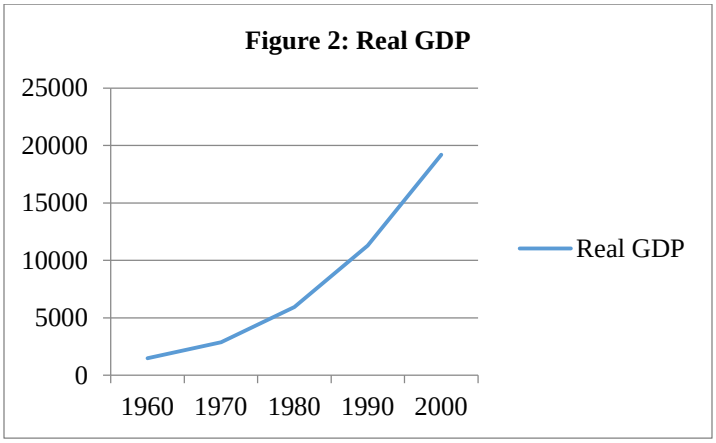
Table 4: Growth in Real GDP

Year	Real GDP
1960	1490.936
1970	2872.189
1980	5944.627
1990	11283.9
2000	19183.93

Source: DataMarket (2014)

Analyzing the data in Table 4, over the years there has been a growth in the real GDP. 1960 was the starting year of Taiwan's experience of improving its economy. The data shows that Taiwan's economic movement was making improvements to the country's situation.





To give a better understanding of the Taiwan's improvements, graphs were made from the previous data. There is a correlation found between the years of schooling and the growth of GDP. Due to the education reform, most of the population became more educated. Taiwan continuously pushed for quality education to ensure that their human capital receives significant positive developments. Considering the human capital theory, when there is an investment in the people of the society then it will lead to enhanced productivity and will influence the society as a whole to strive for a better future (Chaung & Lai). Taiwan's education reform is seen as a success story from its economic development from struggling in an agricultural state to now being a more industrialized and education-based country (Li, 1999).

COMPARING SOUTH AFRICA AND TAIWAN

After examining these two regions, there have been notable similarities and differences. Both regions started out as LDCs, although South Africa is still considered an LDC while Taiwan surpassed this status and became an NIC. Both countries also have implemented a compulsory education that mandates children from ages 5-16 to pursue an education. However, although this is implemented in South Africa, millions of children actually do not continue or even start their education because of the major challenges they face with their poverty state. It is also evident that both governments have been involved in promoting the education systems of their countries, but only one has proven to be successful. Because of Taiwan's strong

government push for education, they have succeeded in growing their nation. However, South Africa has stayed in a low state because there are other stronger environmental factors that prevent them from overcoming their current conditions.

Besides these factors, there is also a different outlook on education as a whole itself. As commonly known by most, there is a strong Asian preference towards education. According to Rodriguez (2011) "Eastern implications of hard work, obligations, and educational success is capitalized and heavily emphasized in Asian philosophies" (para. 2). It is one of the stereotypes that Asians are studious and smart, and in most cases it is so because Asians value education to move higher in their careers to have better lives. However, for the case of South Africa, it is not that they do not want an education, but because of the harsh environment and the very poor economy of most, education is just not a priority for many. The way of life is hard for most Africans and pursuing education is just something they cannot afford. Moreover, there is also growing and continuous civil unrest within the nation of South Africa. Although they have moved on from apartheid, the recent death of Nelson Mandela has now brought on new problems. According to Harding (2013), "Recent headlines have highlighted violent industrial action, the massacre at the Marikana mine, the death of a man dragged behind a police van and the enduringly high crime statistics" (para. 6). So other than their already poverty-stricken land, there are also high crime rates and an unstable government. With all these factors, it is evident why Africans do not put education as one of their main priorities.

CONCLUSION

When looking at these two countries the Big Push theory by Hirschman is considered in this situation. The Big Push theory discusses how poor countries fall into a "poverty trap" in which the country remains stagnant at low levels of growth. In order to become economically better there needs to be a substantial "push" for the economy to take off from its trap and into sustained growth (Walton, 2011). Government intervention plays an important role in contributing the necessary push that the country needs. It is great that Taiwan's government pushed towards the betterment of their society and was able to overcome and progress their nation by building a superior education system. Unfortunately, we see a different story with South Africa. There is a lack of government support in South Africa along with its unstable economic conditions. Moreover, the issue with this is that the Africans have a more difficult and challenging issue to deal with before they can focus their attention on promoting education. When these issues are resolved, only then can South Africa really move forward.

REFERENCES

Chaung, Y & Wei, L. Heterogeneity, Comparative Advantage, and Return to Education: The Case of Taiwan. Retrieved on March 11, 2014 from https://www.gate.cnrs.fr/IMG/pdf/Heterogenous_return_to_education_E1005.pdf

DataMarket. (February 1, 2014). Real GDP per capita. Retrieved on March 10, 2014 from http://datamarket.com/en/data/list/?q=provider:gapminder.

Education Statistics in South Africa 2012 (January 2014). *Department: Basic Education, Republic of South Africa.* Retrieved on March 10, 2014 from http://www.education.gov.za/EMIS/StatisticalPublications/tabid/462/Default.aspx.

Epstein, M.J. & Yuthas, K. (Winter 2012). Redefining education in the developing world. *Stanford Social Innovation Review.* Retrieved on March 9, 2014 from http://www.ssireview.org/articles/entry/redefining_education_in_the_developing_world.

Glewwe, P. et al. (September 2011). School resources and educational outcomes in developing countries: a review of literature from 1990 to 2010. Retrieved on April 27, 2014 from http://faculty.apec.umn.edu/pglewwe/documents/Edinput7.pdf.

Godo, Y. (2012). Center for Economic Institutions Working Paper Series. Retrieved on March 13, 2014 from http://cei.ier.hit-u.ac.jp/English/database/documents/wp2011-12.pdf

Harding, A. (April 6, 2013). After Mandela what will happen to South Africa?. *BBC News.* Retrieved on April 30, 2014 from http://www.bbc.com/news/world-africa-22045720.

Higher Education in Taiwan. (July 2012). Retrieved on March 12, 2014 from http://www.edu.tw/FileUpload/1052-14816%5CDocuments/%E9%AB%98%E6%95%99%E7%B0%A1%E4%BB%8B2012-2013(%E6%B2%92%E6%9C%89%E8%A3%81%E5%88%87%E7%B7%9A).pdf.

Kennedy, E. A. (July 21, 2007). Africa's children struggle for education. *USA Today.* Retrieved on April 29, 2014 from http://usatoday30.usatoday.com/news/education/2007-07-21-africaschild_N.htm.

Li, C. (1999). Sweeping Education Reform in Taiwan: From Primary Education to Lifelong Global Learning. Retrieved on March 10, 2014 from http://www.epa.ncnu.edu.tw/forum/%E8%AB%96%E6%96%87/vol2/vol2-1/3-1%20Sweeping%20Educational%20Reform%20in%20Taiwan%20From%20Primary%20Education%20to%20Lifelong%20and%20Global%20Learning(%E6%9D%8E%E6%8C%AF%E6%B8%85).pdf.

Lin, T. (2003). Education, technical progress, and economic growth: the case of Taiwan. Retrieved on April 29, 2014 from http://www.csus.edu/indiv/l/langd/Lin.pdf.

Mao, Y & Schive, C. Agricultural and Industrial Development in Taiwan. Retrieved on March 10, 2014 from http://citeseerx.ist.psu.edu/viewdoc/download?rep=rep1&type=pdf&doi=10.1.1.139.5132.

Meier, G.M. & Rauch, J.E. (2005). *Leading issues in economic development* (8th ed.). Oxford, NY: Oxford University Press.

Mellor, J. (1995). Agriculture on the Road to Industrialization. Retrieved on March 10, 2014 from http://www.ifpri.org/sites/default/files/publications/mellor95.pdf.

Reviews for National Policies for Education: South Africa (2008). *Organisation for economic co-operation and development (OECD).* Retrieved on April 28, 2014 from http://www.education.gov.za/LinkClick.aspx?fileticket=sKsxhYorWOk%3D&tabid=452&mid=1034.

Rodriguez, A. (November 18, 2011). A debate on the Asian educational mentality. *Ezine Articles.* Retrieved on April 29, 2014 from http://ezinearticles.com/?A-Debate-on-the-Asian-Educational-Mentality&id=6703198.

Walton, M. (2011). The big push: what does it mean, and does it make sense for Ethiopia?. Retrieved on March 13, 2014 from http://www.michaelwalton.info/wp-content/uploads/2011/08/Big-Push-Ethiopia.pdf.

REGIONAL DIFFERENCES IN THE EVOLUTION OF THE U.S. FREEZER INDUSTRY

Emma Bojinova, Canisius College

ABSTRACT

The regional dimension of industry dynamics has emerged as an important research topic in the recent years. Regions within a country differ in terms of availability of natural resources, infrastructure, institutions, population, and labor force with a specific set of skills. Consequently, there can be substantial disparities in prosperity and industrial development across regions.

This paper analyzes the patterns of regional development in the U.S. freezer industry for the period 1946-1981. The analysis is based on the U.S. Census Division of the United States into nine distinct regions/divisions. Along with many other U.S. manufacturing industries, the freezer industry experienced a severe shakeout or a substantial drop in the number of producers over a relatively short period of time. The industry was also highly geographically concentrated in a few regions during the entire study period.

The evolution of freezers is divided into three stages – growth, shakeout, and stabilization. Various statistics are computed to determine whether the different regions exhibited dissimilar dynamics. A survival analysis is used to make further comparisons of the selected regions. Overall, the paper presents some evidence for the dissimilar development of the regions and provides a short discussion about the possible reasons for these differences.

INTRODUCTION

Many industries experience substantial changes in the number of producers over time. A number of studies documented a common evolutionary pattern in different manufacturing industries in the United States (Gort & Klepper, 1982; Klepper & Grady, 1990; Agarwal, 1998) and in the United Kingdom (Simons, 2005). The number of producers initially increased after the inception of a product, reached a peak, and then, suddenly declined. This decline can be gradual or very rapid and is referred to as a shakeout. Products that underwent many and/or substantial technological changes were subject to more severe shakeouts–the number of firms dropped by more than 60% over a short period of time (Agarwal, 1998). Such substantial changes in the number of competing manufacturers significantly affected the market structure in those industries and many of them eventually became oligopolistic in nature.

These common patterns sparked a great interest in researchers to determine the causes of shakeouts. A number of theories were developed to explain shakeouts and the role of technological developments was identified as the main factor influencing competition and the number of firms in an industry (see for example, Utterback & Suàrez, 1993; Jovanovic & MacDonald, 1994; Klepper, 1996, 2001). In short, firms (both incumbents and new entrants) that are unable to develop and/or adopt the new technological innovations become less competitive,

lose some of their customers, experience continuous declines in their profits, and are eventually forced to exit the industry. This triggers a shakeout.

Many of the empirical studies regarding industry shakeouts were conducted at a country level and investigated how the number of producers in a given industry changed over time. They rarely considered the impact of geographic location (for a detailed discussion see Boschma & Lambooy, 1999; Boschma & Frenken, 2011). Recently, researchers started to pay more attention to the locational aspect of industrial dynamics. Regions within a country are often endowed with different factors of productions. They differ in terms of their availability of natural resources, population, human and physical capital, infrastructure, and institutions. Consequently, there can be substantial disparities in development, income distribution, and prospects for economic growth across the different regions.

This paper investigates whether there are regional differences in the evolution of the U.S. home and farm freezer industry. The industry experienced a severe shakeout several years after the mass production of freezers began. This industry was also geographically concentrated in a few regions throughout the study period. The freezer industry was chosen because it has been rarely considered in the industrial studies of shakeouts, but most importantly the regional component of its evolution has not been studied so far. The main research questions of this paper are:

1. Do the evolutionary patterns for the different regions fit the country-level pattern?
2. Can the differences between patterns of regional development/evolution be explained by theories of industry shakeouts or the theories in the regional science literature?

To answer these questions, a historical dataset of firm location is created and analyzed. The source of data is the Thomas' Register of American Manufacturers, which tracks manufacturing establishments over time and provides information about their location (incl. specific address) and approximate capital rating designed by the publisher. The analysis is based on the U.S. Census division of the United States into nine distinct regions. Three of the regions attracted the majority of freezer producers and are considered separately, whereas the other six regions had a relatively small number of firms throughout the years and are combined into one larger region. The evolution of the industry is divided into three stages following Klepper and Grady (1990): growth, shakeout, and stabilization. Various statistics are computed and compared across the different regions during each stage of development, such as the change in the number of establishments, entry and exit rates, proportion of large establishments, percentage drop during the shakeout, etc. The paper provides evidence that the different regions seem to exhibit dissimilar dynamics. Two hypotheses are tested to determine which theories of industry shakeouts or spatial concentration fit better with the patterns of development in the selected industry.

The paper is structured as follows. In section 2, a brief literature review is presented focusing on the main theories of industry shakeouts and the regional science theories. Section 3 provides description of the data, as well as the employed methodology. Section 4 analyzes the industrial dynamics in the different regions. Section 5 provides a short discussion, summarizes the findings, and concludes.

BRIEF LITERATURE REVIEW

The three leading and most relevant theories of industry shakeouts are the "dominant design" by Utterback and Suàrez (1993), the "innovative gamble" of Jovanovic and MacDonald (1994), and Klepper (1996)'s theory of increasing returns to R&D. All of them explain shakeouts as a consequence of some type(s) of technological change. It is difficult to test these theories directly because it is hard to determine the impact of technological change (Klepper & Simons, 2005). However, analysis of entry and exit rates, as well as, of survival rates can provide evidence for which one of these three theories fits more with the available data for the home and farm freezer industry. These theories have different predictions for industrial dynamics with regards to entry and exit patterns and survival of incumbents compared to later entrants.

According to Utterback and Suàrez (1993), the occurrence of a dominant design affects competition and the number of firms in an industry. This dominant design can be one or two specific features of a product or a process innovation that become(s) a standard. Being superior to the other existing design structures it provides an incentive for innovation. Firms that are unable to innovate are forced to exit, which triggers a shakeout in the industry. Only the most capable innovators survive. Exit rates decrease and the number of firms stabilizes after the shakeout. The theory predicts higher survival rates for early entrants or incumbents in the beginning. However, as the least capable innovators exit the market, the survival rates for both early entrants and firms that entered during the shakeout (late entrants) become more similar.

In Jovanovic and MacDonald (1994), shakeouts are caused by a single major refinement following a basic innovation, which increases the optimal size of firms. In each stage firms decide whether to innovate or not and because of that this model is referred to as the "innovative gamble" in the literature. It predicts that the number of firms in an industry first rises due to invention, stabilizes until the refinement, then increases further. As unsuccessful innovators leave the industry, contributing to a shakeout, the number of firms stabilizes. This theory has similar predictions– entry rates drop with the onset of the shakeout, exit rates increase during the shakeout, and eventually the number of firms levels off. Early entrants have a competitive advantage over post-shakeout entrants initially but as late entrants gain experience the difference between survival rates of the pre- and post-shakeout entrants becomes insignificant.

Klepper's (1996) theory predicts that shakeouts are a consequence of a broader evolutionary process arising from increasing returns to scale to R&D. Incumbents benefit more from R&D (especially process innovations) and become more competitive compared to later entrants. Process R&D lowers average cost of production and bigger firms gain advantage due to increasing returns to scale. The decrease in the number of new entrants, along with the increase in the number of firms that exit the industry due to unsuccessful innovations, contribute to a shakeout. Only the most capable entrants are able to compete with the incumbents. The advantage of early entrants over post-shakeout entrants can become even stronger over time which may prevent future entry in this industry. The predictions of this theory are different than those of the "dominant design" and "innovative gamble" - exit rates will not diminish eventually and survival rates of incumbents and post-shakeout entrants will not become similar over time.

Earlier entrants will still have advantage over late entrants and the industry will become even more dominated by some of the incumbents.

The three theories of industry shakeouts do not make any predictions regarding patterns of industrial development in the different regions within a country. In contrast, there is a broad literature focusing on geographic agglomeration and changes in location over time. However, these theories do not specifically address the shakeout phenomenon. Some of these theories are relevant to the scope of this paper and will be briefly discussed next.

Industrial activities may cluster in specific geographical areas due to positive externalities associated with labor market pooling (availability of skilled labor), access to suppliers of specialized equipment and inputs (supply-side factors), proximity to customers (demand-side factors), transportation cost savings, and knowledge spillovers within or across industries (Krugman, 1991). In addition, some regions can grow faster than others based on their ability to generate new knowledge (Romer, 1986).

A number of theories have been developed to explain the possible changes in spatial concentration of industries over time. According to the profit-cycle theory, production activities are likely to be more concentrated initially, but as industries mature they will tend to be more dispersed due to changes in profit opportunities (Markusen, 1985). Some geographical areas can act as "nursery" centers in promoting innovation and will attract younger firms (Duranton & Puga, 2001). These firms will have access to different production technologies that will allow them to experiment and innovate until they find their "ideal" production process. Then, they will move away from the "nursery" centers to places with lower costs of production. The decline in the concentration of firms in a particular geographical area over time can also be due to the diminishing importance of external economies of scale after a product is developed and goes through its growth stage. In the mature phase, the product becomes standardized and specialized skills become less important (Hansen, 1988). This destabilizes the process of local concentration.

On the other hand, clusters may form in specific locations simply due to historical accidents (Krugman, 1991). As firms concentrate in these locations they benefit from being close to their competitors and attract more suppliers of their inputs and workers with specific skills. This leads to a cumulative process that promotes industrial growth and development in these locations and economic activities can continue to stay concentrated in this location for the entire product life cycle. Spinoffs (startups found by at least one person who was previously employed by another producer in the same industry) and spinoff dynamics have a significant role in fostering innovation and regional economic development. For instance, in the U.S. automobile, tire, semiconductors, footwear, high technology, and fashion industries clusters emerged by historical precedents and spinoff dynamics lead to a lock-in effect (Sorenson & Audia, 2000; Moore & Davis, 2004; Neck, Meyer, Cohen, & Corbett, 2004; Klepper, 2007, 2010; Wenting, 2008; Buenstorf & Klepper, 2009, 2010).

The theories of clustering and locational change over the product life cycle predict different patterns of industry dynamics. In the first case, regional concentration of production activities will tend to be relatively constant over time. In the second case, the concentration of an industry will decline during the later stages of a product development. One of the goals of this paper is to check which theory fits better with the data for the U.S. freezer industry.

DATA DESCRIPTION AND METHODOLOGY

The data source for the following analysis is the Thomas' Register of American Manufacturers. The data were compiled from microfilms for the period 1946-1981. The Register is a relatively reliable source of data and has been typically used in the industry shakeouts literature. It contains information about the name and address (state, city, and often detailed address) of the U.S. manufacturers of home and farm freezers. It also provides a capital rating (a non-numeric coding) that can be used as a crude measure of firm size. Establishments with capital above $1,000,000 have a rating of AAAA, establishments with capital above $500,000 but below $1,000,000 have a rating of AAA, establishments with capital between $200,000 and $499,999 receive a rating of AA, etc. Establishments with unknown capital are given a rating of X. Producers with production facilities in different cities are counted as separate establishments.

The home and farm freezer industry was selected for two reasons - first, it experienced a severe shakeout in 1955, and second, it had enough observations for conducting the geographical analysis. The time span is limited to 1981 because the main focus of this paper is on the shakeout period and this period ended several years prior to 1981. The list of freezer producers was slightly adjusted for minor gaps in the time series data.

The regional analysis is based on the U.S. Census partition of the United States into 9 distinct divisions or regions: New England (Region 1), Middle Atlantic (Region 2), East North Central (Region 3), West North Central (Region 4), South Atlantic (Region 5), East South Central (Region 6), West South Central (Region 7), Mountain (Region 8), and Pacific (Region 9). Regions 2, 3, and 4 had the largest concentration of freezer producers throughout the study period. In contrast, regions 1, 5, 6, 7, 8, and 9 had just a few establishments over the years (at most 3 in a given year) and were combined into a single region named Region 1, 5-9.

The paper utilized the algorithm developed in Klepper and Grady (1990) to determine the different stages of industrial evolution for each region described above and for the industry as a whole. In particular, the life cycle of a product consists of 3 stages. Stage 1 is the growth period – it starts with the product inception and ends at the peak year (i.e., the year with the maximum number of producers). When there is more than one year with the same maximum number of establishments, the peak year is the one that has the highest average number of establishments in the following three years. Stage 2 is the shakeout period. The number of establishments declines during this stage. The shakeout period ends in the year after which the annual change in the number of establishments averaged over the next 5, 10, 15, 20, ... years is greater than -1.0% of the peak number of establishments. Stage 3 is the post-shakeout period.

REGIONAL ANALYSIS

The missing component of the industry evolution studies conducted at the country level is the geographic dimension of industry dynamics. Different regions often differ in terms of their endowments, access to resources and institutions, as well as the level of competition among firms. Because of that, it is likely to expect different patterns of industrial evolution for the different regions. These disparities in regional development and their causes constitute an

important research question because they affect the rates of local economic growth, income distribution, and the well-being of local residents.

In this section, the regional development of the U.S. freezer industry will be analyzed by providing a graphical presentation of the different regions in terms of entry, exit, and number of establishments over time and by comparing a number of statistical measures for each stage of development and also the survival rates of pre- and post-shakeout entrants across regions and in the entire industry. In addition, the implications of the three theories of shakeouts and the theories of regional concentration will be considered in order to determine which theory fits better with the data available for the freezer industry.

Home and farm freezer units were introduced into the U.S. market in 1929 (see for example Gort & Klepper, 1982; Klepper & Grady, 1990). Their production was suspended during the Great Depression and the World War II. After the end of the war freezers became very popular and their production was resumed. Initially, freezers were built as a compartment in a larger refrigerator, but they were expensive to run and required thawing to remove the build-up of ice around the freezing units. The technology of frost-free refrigeration was developed in 1950s, which made freezers more efficient and easier to maintain (Ormrod, 1990). The design and energy efficiency of freezers were improved multiple times over the next three decades. Many other advances such as automatic ice units and self-compartmentalized freezing units were also developed.

The major freezer producers in the 1940s were Amana Society (later acquired by Raytheon), Maytag, General Electric, General Motor's Frigidaire, and Westinghouse Electric. Amana Society and Maytag were established in Iowa, General Electric in Connecticut, and the last two companies operated in Ohio. The majority of freezer manufacturers were located in Region 3 (East North Central) and Region 2 (Middle Atlantic), there were some in Region 4 (West North Central), and just a few in the other regions. Region 3 can be considered as the industry center because it had the highest concentration of establishments throughout the years.

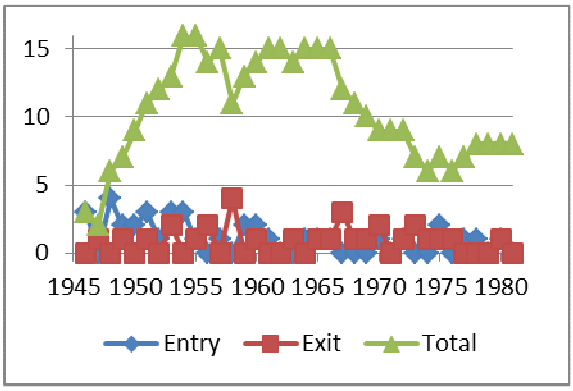
Figure 1 presents the total number of establishments, entry and exit counts by region in the freezer industry for the period 1946-1981. The four regions appear to have experienced different patterns of development. The declines in the number of establishments in Regions 2, 3, and 4 seem to be more substantial compared to the combined region (Regions 1, 5-9). In addition, the decline in the latter region started later than in the other three regions. Moreover, a comparison of the regional patterns with the entire industry shows that these patterns of industry dynamics look dissimilar. Perhaps, just the pattern for Region 3 seems to fit to some extend with the one for the industry as a whole. Entry and exit patterns also appear to be different across the different regions. Additional statistics are necessary to make better comparisons among the regions and the entire industry.

As described in the previous section, the evolution of freezer industry is divided into three stages: growth, shakeout, and stabilization following Klepper and Grady (1990). A number of important descriptive statistics are computed for each of the regions and the entire industry by stage of development. These statistics are summarized in Table 1.

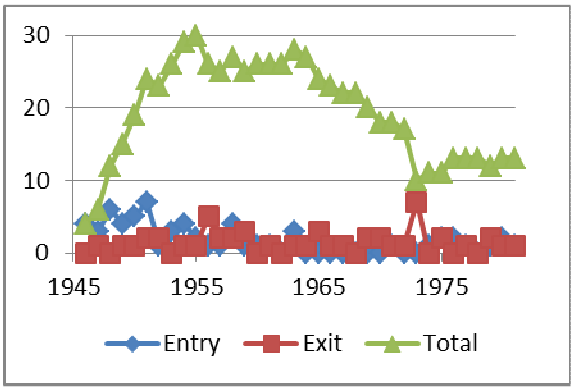
The average net entry rate is computed by averaging the net entry rates in each stage (i.e., the difference between the number of establishments entered in and those that left the region/industry in a given year divided by the total number of establishments in the previous year). This statistic is useful in comparing the entry and exit patterns across regions. Similarly,

average net entry rates for the first quarter of the shakeout period for the entire industry (not for a particular region) is calculated for all regions. For instance, the shakeout in the freezer industry began after 1955, the peak year, and ended in 1973. The duration of the shakeout was 18 years. One quarter of 18 years is 4.5 years. Therefore, years 1956-1960 were considered as the first quarter of the shakeout.

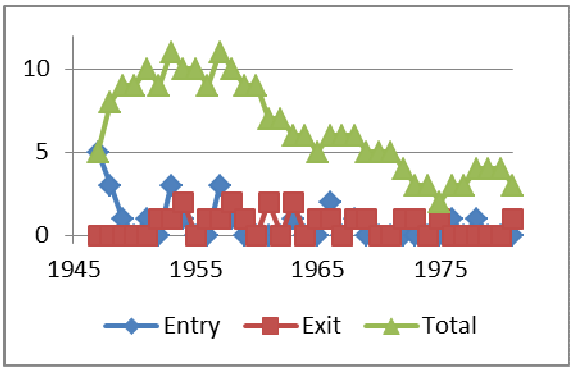
Figure 1: Evolution of Freezer Industry by Region and for the Entire Industry



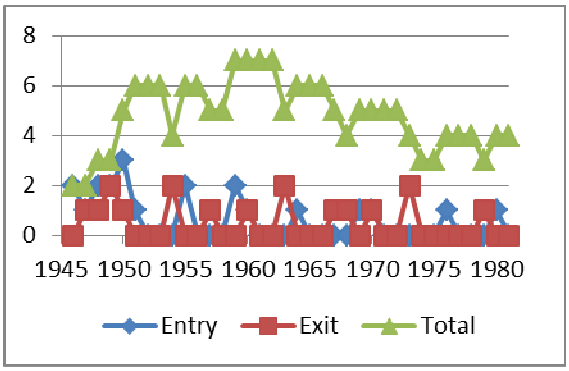
Panel a: Region 2



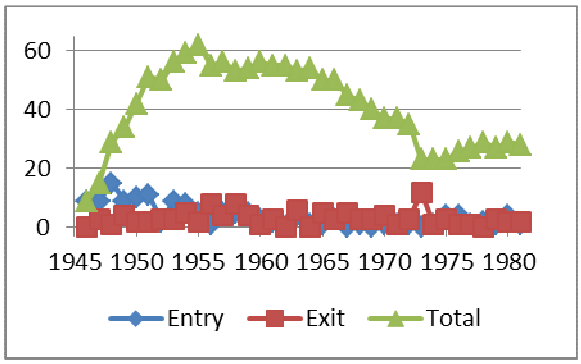
Panel b: Region 3



Panel c: Region 4



Panel d: Regions 1, 5-9



Panel e: All Regions (Entire Industry)

Several studies found a positive relationship between firm size and process innovations (Mansfield, 1981; Scherer, 1991; Cohen & Klepper, 1996), which in turn affect firm survival. The capital rating is utilized to distinguish freezer producers by their size. The percentage of

large establishments is computed as the ratio of the number of distinct manufacturing establishments with a capital rating of AAA or higher to the total number of establishments in each stage and region. All establishments that entered or exited are counted.

The age of establishments can be used as a proxy for their production experience (productivity can increase over time due to the "learning by doing" effect) or innovative capabilities. Incumbents were found to have substantially lower hazard rates compared to later entrants and this outcome was related to the higher probability of survival of innovators and the higher likelihood of early entrants to innovate (Klepper & Simons, 2005).

The percentage drop in stage 2 and percentage drop during the first quarter of the shakeout for the entire industry are two other measures listed in Table 1. They are useful in determining the severity of the shakeout in each region and the entire industry. The first measure is computed as the difference between the peak number of establishments and the number of establishments at the end of the shakeout divided by the peak number. The percentage drop during the first quarter of the shakeout is calculated in a similar way but in this case we subtract the number of establishments at the end of the first quarter.

Table 1
Descriptive Statistics by Stage and Region

	Region 2	Region 3	Region 4	Regions 1, 5-9	All Regions
Stage 1 (Growth)	(1946-1954)	(1946-1955)	(1946-1953)	(1946-1959)	(1946-1955)
Average net entry rate	0.34	0.28	0.16	0.14	0.27
% large establishments	36.36	52.63	53.85	50.00	49.44
Average age	2.76	3.35	3.69	3.38	3.68
Peak number of establishments	16	30	11	7	62
Stage 2 (Shakeout)	(1955-1973)	(1956-1973)	(1954-1974)	(1960-1973)	(1956-1973)
Average net entry rate	-0.036	-0.052	-0.053	-0.043	-0.049
Average net entry rate in 1st Q	-0.012	-0.025	-0.014	0.058	-0.018
Percentage drop in stage 2	56.25	66.67	72.73	42.86	62.90
Percentage drop in 1st Q	12.50	13.33	18.18	-16.67	9.68
% large establishments	40.00	68.09	61.11	40.00	58.51
Average age	9.93	12.40	9.06	14.8	11.90
Stage 3 (Stabilization)	(1974-1981)	(1974-1981)	(1975-1981)	(1974-1981)	(1974-1981)
Average net entry rate	0.024	0.036	0.036	0.052	0.027
% large establishments	81.82	85.71	75.00	80.00	80.49
Average age	12.09	14.35	12.50	19.6	14.50
Number of establishments in 1981	8	13	3	4	28

A comparison of the different regions reveals the following patterns. The duration of stage 1, or the period of growth, was slightly shorter in regions 2 and 4, but it was four years longer in the combined region as compared to region 3 and the entire industry. In other words, the shakeout started earlier in regions 2 and 4 and later in regions 1, 5-9. A possible explanation for this outcome is that technological changes peaked faster in regions 2, 3, and 4 than in regions

1, 5-9. The average age of establishments located in region 4 was the highest among all regions, which suggests that freezer producers in this region had on average more experience compared to their competitors in the rest of the country. In addition, this region had the highest proportion of large establishments and one of the lowest average net entry rates indicating that the growth period was relatively rapid and involved more large rather than small establishments. Regions 2, on the other hand, had both lower average age and lower percentage of large establishments. This suggests that stage 1 in region 2 was characterized by a larger number of small and younger establishments, which is also evident by the highest average net entry rate in this region compared to the other regions. Region 3, the industry center, developed in a similar way as the entire industry.

The shakeout lasted longer in region 4 compared to regions 3 and 2 (three and two years longer, respectively). The shortest span of the shakeout period was observed for regions 1, 5-9. The average net entry rates were negative indicating that exit rates exceeded entry rates in all regions during stage 2. The percentage drop in the number of establishments exceeded 55% in regions 2, 3, and 4. Only the combined region (1, 5-9) had a decline that was lower than 50%, which indicates that the shakeout was not so severe in this location. The statistics for the first quarter of the shakeout confirm this finding–the average net entry rate was positive and the number of establishments did not drop, it actually went up by 16.67%. This also is consistent with the graphical depiction in figure 1. An interesting observation is that the combined region had the highest average age in stage 2. This is due to the fact that some of the youngest establishments did not survive during the shakeout.

Region 4 was affected most in comparison to the other regions. It had the lowest average net entry rate and the highest percentage drop in stage 2 and during the first quarter of the shakeout for the entire industry. The average age of all establishments in this region was the lowest of all regions signifying that some of the earlier entrants also ended up leaving the industry.

Region 3, which had the highest concentration of freezer producers, also experienced one of the biggest drops in the number of establishments, both during the first quarter and for the whole shakeout period. However, it had the biggest percentage of large establishments and one of the highest average ages in stage 2. This suggests that, even though the shakeout was more severe in this region, it affected mainly the late entrants and the smaller establishments. It is consistent with the findings in Boschma and Wenting (2007) that the concentration of many automobile producers in a given region in Great Britain negatively affected the survival rates of new entrants and this impact was stronger during the later stages of industry evolution.

In stage 3, all regions were characterized by a larger number of bigger and older or more experienced establishments. The average age was typically above 12 years and the percentage of large establishments was at least 75 in each region. This confirms the results found in Cefis and Marsili (2006) that small and young firms are at a higher risk of exit. It is also supported by Klepper and Simons (2000) finding that larger firms had in general higher survival rates in the U.S. tire industry. The average net entry rate was relatively low, with just a few entries and exits per year. Region 3 had the highest proportion of large establishments, one of the highest average ages, and the biggest number of operating manufacturing establishments in 1981.

To determine which of the theories of industry shakeouts and of clustering, discussed in the literature review section, fit more with the regional patterns of development in the freezer industry the following two hypotheses will be tested:

H1 *If the shakeout in the freezer industry was a consequence of some kind of technological change(s) as explained by the "dominant design" or "innovative gamble" theories, the survival rates of post-shakeout entrants should be lower initially, but over time, as those firms gain experience, the differences in survival rates of early and late entrants should become more and more similar. In contrast, if the shakeout was caused by a broader evolutionary process arising from increasing returns to scale to R&D (Klepper, 1996) rather than a particular technological development, the difference in survival rates of early and post-shakeout entrants will not become smaller and smaller over time.*

H2 *If spatial concentration changes over time, manufacturing activities should initially be concentrated in a few core regions or industry centers, but as profit opportunities change (Markusen, 1985), when the "ideal" production process is found (Duranton & Puga, 2001), or when the product becomes standardized and firms begin to search for cost savings (Hansen, 1988), more establishments will move their production facilities to lower-cost locations. This will lead to a decline in the concentration of firms in the core region(s). However, if regional concentration is due to historical accidents and spinoff dynamics led to a lock-in effect, the concentration of producers should not change significantly when an industry goes through the different stages of development.*

Table 2 lists the survival rates of establishments in each region by their time of entry (i.e., before or after the beginning of the shakeout). These survival rates represent the percentage of establishments located in a given region that survived at least five, ten, fifteen, twenty, or twenty-five years. The survival rates of pre-shakeout entrants were typically higher than the rates of post-shakeout entrants in all regions (with one exception) and also for the entire industry as measured by the percentage of establishments that survived 5 or more years, or at least ten years. This is consistent with the predictions of the three theories of industry shakeouts.

Table 2
Survival Rates of Pre- and Post-Shakeout Entrants by Region

	Survived at least 5 years (%)	Survived at least 10 years (%)	Survived at least 15 years (%)	Survived at least 20 years (%)	Survived at least 25 years (%)
Region 2 pre-shakeout entrants	59.09	45.45	40.91	13.64	13.64
Region 2 post-shakeout entrants	56.25	25.00	6.25	6.25	0.00
Region 3 pre-shakeout	67.50	47.50	42.50	25.00	15.00
Region 3 post-shakeout entrants	59.09	31.82	13.64	9.09	4.55
Region 4 pre-shakeout	71.43	42.86	14.29	7.14	7.14
Region 4 post-shakeout entrants	70.00	20.00	10.00	0.00	0.00
Regions 1, 5-9 pre-shakeout entrants	46.46	40.00	26.67	20.00	20.00
Regions 1, 5-9 post-shakeout entrants	50.00	33.33	0.00	0.00	0.00
All pre-shakeout entrants	63.22	44.83	34.48	18.39	14.94
All post-shakeout entrants	61.54	28.85	9.62	5.77	1.92

However, the long-term survival rates (at least 15, 20, or 25 years) of pre- and post-shakeout entrants did not become similar over time in any of the regions. To determine whether these survival rates were really different the following procedure was used. The standard deviation of the proportion in a Binomial experiment with success probability p is $\sqrt{\frac{p(1-p)}{n}}$. Plugging in the values of p in this formula, one can compute the standard errors in each case. If a difference in proportions is at least two standard deviations greater than any of the standard deviations, this difference will be considered large enough to conclude that survival rates were different. A formal testing for differences in two proportions was inappropriate in most of the cases because *np* or *n(1-p)* was not greater than 5, which violated the assumptions of this test.

Table 3 lists the differences in proportions of pre-and post-shakeout establishments that survived a given number of years (at least 15, 20, or 25) and the associated standard deviations for each region and the entire industry. For instance, considering the proportion of establishments that survived at least 15 years in region 3, the standard deviations for pre- and post-shakeout entrants are 0.078162 (or 7.82%) and 0.073173 (or 7.32%), respectively. Because the difference 42.5% - 13.64% = 28.86% equals more than three times any of the standard errors, the result is obvious that pre-shakeout entrants had higher survival rates than post-shakeout entrants and no formal testing is needed.

Table 3
Differences in Proportions of Survival Rates between Pre- and Post-Shakeout Entrants by Region

	Survived at least 15 years		Survived at least 20 years		Survived at least 25 years	
Region 2: Difference in proportions	0.35	diff.	0.07		0.14	diff.
Standard deviations	0.10	0.06	0.07	0.06	0.07	0.00
Region 3: Difference in proportions	0.29	diff.	0.16	diff.	0.10	
Standard deviations	0.08	0.07	0.07	0.06	0.06	0.04
Region 4: Difference in proportions	0.04		0.07		0.07	
Standard deviations	0.09	0.09	0.07	0.00	0.07	0.00
Regions 1,5-9: Difference in proportions	0.27	diff.	0.20	diff.	0.20	diff.
Standard deviations	0.11	0.00	0.10	0.00	0.10	0.00
All: Difference in proportions	0.25	diff.	0.13	diff.	0.13	diff.
Standard deviations	0.05	0.04	0.04	0.03	0.04	0.02

The differences in proportions of early and late entrants that survived at least 15, 20, or 25 years are typically at least two times greater than the larger of the standard errors in all regions besides region 4. The results for region 4 are inconclusive, but considering the fact that none of the post-shakeout entrants survived more than 20 years in this region, there is some evidence that late entrants were at disadvantage. The small sample sizes make the standard deviations larger and it is really difficult to get significant results.

Overall, the differences in survival rates of pre-and post-shakeout entrants became even larger as the industry matured, which does not support the "dominant design" and "innovative

gamble" theories. Instead, this pattern fits more with the theory of increasing returns to scale in R&D that shakeouts are a consequence of a broader evolutionary process in which technological change leads to increasing returns and gives an advantage to early entrants. This thesis is also supported by the statistics provided in Table 2, which show that the proportion of large and older establishments in all regions is much higher in stage 3 compared to stages 1 and 2.

Table 4 lists the proportions of establishments located in each region by stage of development. It can be used to determine whether concentration of freezer producers in a given location changed over time (hypothesis 2). The fluctuations in the concentration of establishments from stage 1 to stage 2 or from stage 2 to stage 3 in the different regions seem to be relatively small.

Table 4
Proportion of Establishments in Each Region by Stage

	Region 2	Region 3	Region 4	Regions 1, 5-9
Stage 1: Growth	24.719	42.697	14.607	17.978
Stage 2: Shakeout	28.571	44.762	17.143	9.524
Stage 3: Stabilization	26.829	51.220	9.756	12.195

Two different tests were conducted to check whether these changes in concentration were statistically significant. The first one was a test for differences between two proportions. All, but in one case, the test assumptions were satisfied. Those hypothesis tests did not provide evidence that population proportions were different at the 5% level of significance. The second test was a chi-square test for independence. Four contingency tables were created. The test results showed that the number of establishments located in each region was not dependent on the stage of development at the 5% level of significance. In summary, the concentration of establishments in each region in the freezer industry did not change significantly over the product life cycle, which fits more with the theory of concentration due to historical accidents (Krugman, 1991) and the role of spinoffs in creating a lock-in effect.

CONCLUSION

This paper analyzed the patterns of regional evolution in the American freezer industry for the period 1946-1981. All regions went through three stages of industry development - growth, shakeout, and stabilization. This supports the country-level patterns acknowledged by several researchers for the U.S and the United Kingdom.

The graphical depiction showed that the evolutionary patterns for the different regions looked dissimilar, which was also confirmed by the various statistics provided in Table 1. The duration of the growth period was slightly shorter in regions 2 and 4 and longer in the combined region compared to region 3 and the entire industry. The shakeout started earlier in regions 2, 3 and 4. This outcome suggests that technology peaked faster in these regions, and as competition among producers became more intense some of the unsuccessful innovators ended up leaving the industry causing the earlier beginning of the shakeout.

The shakeout was more severe in regions 2, 3, and 4 than in the combined region, where it was much shorter. Regions 3 and 4 had the lowest average net entry rates and the highest percentage drops in stage 2 and also during the first quarter of the shakeout. Region 3, which can be designated as the industry center, had the biggest percentage of large establishments and one of the highest average ages in stage 2 suggesting that the shakeout mainly affected late entrants and smaller establishments. This supports Boschma and Wenting (2007) and Cefis and Marsili (2006) that small and young firms are more likely to exit. It also conforms to the idea that large establishments benefit more from knowledge spillovers, grow even larger, become more competitive, and survive longer.

The survival patterns in all regions were analyzed to determine which theory of industry shakeouts was more consistent with the evolution of the freezer industry. The survival rates of pre-shakeout entrants were generally higher than those of post-shakeout entrants in all regions and in the entire industry throughout the study period. The differences in survival rates did not become smaller and smaller over time, instead, they became larger. This finding does not support the "dominant design" and "innovative gamble" theories, but fits more the predictions of the "increasing returns" theory to R&D that eventually some of the early entrants or incumbents will be dominating the industry.

The freezer industry did not experience significant changes in regional concentration over the product life cycle. Region 3 continued to be the region with the highest proportion of establishments during all stages of industry development, followed by region 2. The absence of changes in spatial concentration is in accordance with the theories of clustering due to spinoff dynamics and the agglomeration forces discussed in Krugman (1991). Manufacturing establishments located in the industry center or in the core regions tend to benefit from knowledge spillovers, labor-market pooling, and proximity to suppliers of inputs.

The unequal development of regions has important consequences for regional growth and prosperity of local residents. Regional industrial policies focusing on innovation and improving local conditions for businesses are likely to attract more manufacturing establishments and the self-reinforcing effects of agglomeration can expand the prospects of economic growth in these regions both in the short and long term.

REFERENCES

Agarwal, R. (1998). Evolutionary trends of industry variables. *International Journal of Industrial Organization,* 16, 511-525.

Boschma, R.A. & K. Frenken (2011). The emerging empirics of evolutionary economic geography. *Journal of Economic Geography*, 11, 295-307.

Boschma, R.A. & J.G. Lambooy (1999). Evolutionary economics and economic geography. *Journal of Evolutionary Economics*, 9, 411-429.

Boschma, R.A. & R. Wenting (2007). The spatial evolution of the British automobile industry: Does location matter? *Industrial and Corporate Change*, 16, 213-238.

Buenstorf, G. & S. Klepper (2009). Heritage and agglomeration: The Akron tyre cluster revisited. *The Economic Journal*, 119, 705-733.

Buenstorf, G. & S. Klepper (2010). Why does entry cluster geographically? Evidence from the U.S. tire industry. *Journal of Urban Economics*, 68, 103-114.

Cefis, E. & O. Marsili (2006). Survivor: The role of innovation in firms' survival. *Research Policy*, 35, 626–641.

Cohen, W.M. & S. Klepper (1996). Firm size and the nature of innovation within industries: The case of process and product R&D. *The Review of Economics and Statistics*, 78, 232-243.

Duranton, G., & D. Puga (2001). Nursery cities: Urban diversity, process innovation, and the life cycle of products. *The American Economic Review,* 91, 1454-1477.

Gort, M. & S. Klepper (1982). Time paths in the diffusion of product innovation. *The Economic Journal*, 92, 630-653.

Hansen, N. (1988). Regional consequences of structural changes in the national and international division of labor. *International Regional Science Review*, 22(2): 121-136.

Jovanovic, B. & G.M. MacDonald (1994). The life cycle of a competitive industry. *The Journal of Political Economy*, 102, 322-347.

Klepper, S. (1996). Entry, exit, growth, and innovation over the product life cycle. *The American Economic Review*, 86, 562-583.

Klepper, S. (2007). Disagreements, spinoffs, and the evolution of Detroit as the capital of the US automobile industry. *Management Science*, 57, 616-631.

Klepper, S. (2010). The origin and growth of industry clusters: The making of Silicon Valley and Detroit. *Journal of Urban Economics*, 67, 15-32.

Klepper, S. & E. Grady (1990). The evolution of new industries and the determinants of market structure. *RAND Journal of Economics*, 21, 27-44.

Klepper, S. & K.L. Simons (1996). Innovation and industry shakeouts. *Business and Economic History*, 25 (1), 81-89.

Klepper, S. & K.L. Simons (2000). The making of an oligopoly: Firm survival and technological change in the evolution of the U.S. tire industry. *The Journal of Political Economy*, 108, 728-760.

Klepper, S. & K.L. Simons (2005). Industry shakeouts and technological change. *International Journal of Industrial Organization*, 23 (1-2), 23-43.

Krugman, P. (1991). Increasing returns and economic geography. *Journal of Political Economy,* 99 (3), 483– 499.

Markusen, A. (1985). *Profit cycles, oligopoly, and regional development*. Cambridge, MA: MIT Press.

Moore, G. & K. Davis (2004). Learning the Silicon Valley way in T. Bresnahan and A. Gambardella (eds.), *Building High-tech Clusters: Silicon Valley and Beyond*, Cambridge University Press: Cambridge, England.

Neck, H.M., Meyer, G.D., Cohen, B. & A.C. Corbett (2004). An entrepreneurial system view of new venture creation. *Journal of Small Business Management*, 42, 190-208.

Romer, P.M. (1986). Increasing returns and long-run growth. *Journal of Political Economy,* 94, 1002– 1037.

Simons, K.L. (2005). Product market characteristics and the industry life cycle. *Cornell University*, Rensselaer.

Sorenson, O. & P.G. Audia (2000). The social structure of entrepreneurial activity: Geographic concentration of footwear production in the United States, 1940-1989. *American Journal of Sociology*, 106, 424-464.

Utterback, J.M. & F.F. Suàrez (1993). Innovation, competition, and industry structure. *Research Policy*, 22, 1-21.

Wenting, R. (2008). Spinoff dynamics and the spatial formation of the fashion design industry. *Journal of Economic Geography*, 8, 593-614.

A CLASSROOM EXERCISE FOR TEACHING THE THEORY OF THE SECOND BEST IN A DUOPOLY SETTING

James Sawler, Mount Saint Vincent University

ABSTRACT

The Theory of the Second Best is an important proposition for students of public policy and economics. It states that, when markets are characterized by imperfections, efforts to enhance welfare by removing other imperfections can adversely reduce economic welfare. This article presents a simple classroom demonstration for teaching the Theory of the Second Best in an oligopoly setting. In the exercise, students represent firms in a duopoly choose a profit-maximizing price under specific market conditions. The results demonstrate that imperfect information and product heterogeneity, two market "imperfections," make cartel coordination challenging and, thus, can enhance market performance. A description of a typical classroom discussion follows.

The effectiveness of the exercise was tested by splitting a principles class into an experimental group and a control group. The results showed that the mean grade of students taught with the classroom demonstration was significantly higher than those taught by more traditional methods.

INTRODUCTION

The Theory of the Second Best (TSB) states that, when markets are characterized by imperfections, efforts to enhance welfare by removing other imperfections can adversely reduce economic welfare. Originally an insight emanating from general equilibrium theory (Lipsey and Lancaster, 1956), the TSB can also apply to partial equilibrium situations. A common example is of a mining monopoly where the output creates pollution -- an external cost on society. Breaking up the monopoly into many firms, an effort to remove one market imperfection could reduce overall welfare as the firms expand production and, therefore, create more pollution.

Though often overlooked, this seemingly paradoxical effect on welfare makes the TSB an important concept for students of economic policy. The narrative common in most textbooks describes how the "ideal" state of perfect competition in a market economy optimizes economic welfare. Deviations from perfect competition are described as reducing economic welfare, requiring government intervention (Ragan, pp. 281-303). However, the TSB reminds students that care must be taken when prescribing policy. Simply removing a market imperfection does not guarantee enhanced welfare and could make the outcome worse. J.M. Clarke (1940) introduced the concept of a remedial imperfection - a deviation from perfect competition that improves the functioning of the market.

This article presents a simple classroom exercise for teaching the TSB to undergraduate students of economic policy. Using a duopoly setting, the exercise acts as an effective starting point for discussions on the concepts of market imperfections and cartel behavior.

CLASSROOM EXERCISES AND DISCUSSION

This exercise is directed towards students who have completed a principles course in microeconomics (discussion questions are presented in italics). The class begins by asking students to recall the criteria of a perfectly competitive market. Collectively, students should have little difficulty remembering most of these, though (significantly for this demonstration) they may need to be reminded about the criterion of complete information. Students can also be reminded of the theoretical benefits of perfect competition towards economic welfare.

Can you name a market that meets all of the criteria for perfect competition? Clearly, this is a challenge. While some markets may be close, students begin to realize that perfect competition is an unattainable ideal.

Consider a market such as the automobile industry; what imperfections exist in this market? How about the airline industry (choose any well-known industries)? Students will recognize that these real markets are characterized by imperfections such as fewness, differentiated products, entry barriers, and incomplete information.

The instructor informs the students that they will now participate in an exercise that will help them understand the difficulties of setting economic policy for imperfect markets. The students are grouped into pairs and told they will each represent a hypothetical firm in a duopoly. They will be presented with different scenarios; their objective in each is to maximize their individual firm's profits.

Scenario 1

The first student, representing firm A, receives a cue card with the following information:

Firm A

Homogeneous Product
Market Demand: $Q = 1000 - 10p$
Your Firm's Constant Marginal and Average Cost: $MC_A = ATC_A = \$15$
Set your price to maximize your profits. No cooperation or communication.

The second student, representing firm B, receives the following information:

Firm B

Homogeneous Product
Market Demand: $Q = 1000 - 10p$
Your Firm's Constant Marginal and Average Cost: $MC_B = ATC_B = \$10$
Set your price to maximize your profits. No cooperation or communication.

Thus, this is a Bertrand duopoly situation with a homogeneous product and constant marginal costs. The only difference between the two firms is that Firm B has a cost advantage. Before the students begin, it is useful to ask: what imperfections characterize this market? Most students will identify that, with only two firms and no cooperation, the market is characterized by fewness and incomplete information. Depending on the level of the students, it may also be helpful to remind them that, with a homogeneous product, the competitor with the lower price will capture all of the demand.

Students typically realize that they must set their price as low as it is possible to still make a profit. Thus, firm A's price is typically set just above \$15, while firm B's is set just above \$10. So, Firm B captures all of the demand. The instructor asks the students to calculate the level of output and the consumer surplus for the prices they set. Depending on the student level, the instructor can assist with a numerical example.

Scenario 2

The first student, representing firm A, receives the following information:

Firm A

Homogeneous Product
Market Demand: $Q = 1000 - 10p$
Your Firm's Constant Marginal and Average Cost: $MC_A = ATC_A = \$15$
Your Competitor's Constant Marginal and Average Cost: $MC_B = ATC_B = \$10$
Set your price to maximize your profits. No cooperation or communication.

The second student, representing firm B, receives the following information:

Firm B

Homogeneous Product
Market Demand: $Q = 1000 - 10p$
Your Firm's Constant Marginal and Average Cost: $MC_B = ATC_B = \$10$
Your Competitor's Constant Marginal and Average Cost: $MC_A = ATC_A = \$15$
Set your price to maximize your profits. No cooperation or communication.

The only difference from Scenario 1 is that the firms now know the costs facing their competitor.

Most of the students playing Firm B now recognize that they can raise their price to just below \$15 and still capture all of the demand.

Discussion

Compared to Scenario 1, what market imperfection was removed in Scenario 2? Students typically recognize that information became more complete.

How did the prices, outputs, and consumer surpluses differ between the two scenarios? The students have calculated that price is higher and both output and consumer surplus are lower in Scenario 2.

So, when we removed a market imperfection by making information more complete, consumer welfare decreased. Given that perfect competition maximizes consumer welfare, wouldn't you expect the removal of an imperfection to improve consumer welfare? At this stage, students begin to grasp the supposed paradox that is the TSB, which the instructor can formally define.

If the instructor only wants a quick demonstration of the TSB, she may wish to stop here and move on to a discussion of the TSB's implications for setting economic policy (see below). If she wants another demonstration of the TSB that provides insights into the nature of cartels, she can continue with scenarios 3 and 4.

Scenario 3

The first student, representing firm A, receives the following information:

Firm A

Differentiated Product
Your Firm's Demand: $Q_A = 500 - 7p_A + 3p_B$
Your Firm's Constant Marginal and Average Cost: $MC_A = ATC_A = 10$
Communicate with your competitor and set a price to maximize your profits.

The second student, representing firm B, receives the following information:

Firm B

Differentiated Product
Your Firm's Demand: $Q_B = 500 - 8p_B + 2p_A$
Your Firm's Constant Marginal and Average Cost: $MC_B = ATC_B = 10$
Communicate with your competitor and set a price to maximize your profits.

Note that the two firms' demands are such that if $p_A = p_B$, then the market demand could be represented by $(Q_A + Q_B) = Q = 1000 - 10p$. This makes scenario 3 roughly comparable to the others. The instructor can point this out to the students.

What are the market imperfections inherent in this scenario? The students typically recognize product differentiation as an "imperfection." Some in the class also recognize that communication sets up the possibility of a cartel. If no students raise this possibility, the instructor should do so.

The results for scenario 3 are varied. Some students cooperate and set a price through trial and error. Others set their price independently. Typically, whether they coordinate or not, students use the ATC of $10 as an anchor and add a substantial markup, so that prices are generally in the range between $15 and $25. The instructor may wish to ask students how they set their prices. This can invoke a discussion of price setting in real markets and the debate between rational price-setting and administered pricing.

As above, the instructor asks the students to calculate the joint output that results from their prices. While calculating consumer surplus is a challenge in this setting, the instructor can remind students that lower prices and greater output typically result in greater consumer welfare.

Scenario 4

Firm A

Homogeneous Product
Market Demand: Q = 1000 - 10p
If you and your competitor set the same price, you will share the market equally.
Your Firm's Constant Marginal and Average Cost: $MC_A = ATC_A = 10$
Communicate with your competitor and set a price to maximize your profits.

Firm B

Homogeneous Product
Market Demand: Q = 1000 - 10p
If you and your competitor set the same price, you will share the market equally.
Your Firm's Constant Marginal and Average Cost: $MC_B = ATC_B = 10$
Communicate with your competitor and set a price to maximize your profits.

This scenario returns to a homogeneous product. Communication between the firms increases the likelihood of collusion.

The great majority of students decide to collude with their competitors and set prices considerably higher than any of the previous scenarios. Depending on the level of the students, some are able to calculate the price and quantity that maximizes joint profits (p = $55, Q = 90).

Discussion

How do scenarios 3 and 4 contrast with respect to your ability to coordinate and set prices? Most students find that the complexity added by product differentiation in scenario 3 makes setting prices much more difficult.

How do price and quantity for scenario 4 compare with scenario 3? For almost all students, price is higher and quantity lower for scenario 4. The homogeneous product has made it much easier for them to coordinate prices. The instructor points out that this is another example of the TSB -- when the market imperfection of a differentiated product is removed, collusion becomes easier and the resulting higher prices and lower output decreases consumer welfare.

So, we have seen two examples of the TSB in an oligopoly setting (the instructor may wish to provide other examples such as the mining monopoly described above). *What are the implications for economic policy?* The principal insight is that policy-makers must proceed with caution; they cannot simply assume that removing market imperfections will improve economic welfare (although often it may). Such policy measures may have unexpected negative consequences. The implications of any policy intervention should be examined on a individual basis.

Cartel Formation

The instructor may wish to continue with a discussion on the factors conducive to forming cartels. *Contrasting scenario 4 with the others, what are the characteristics that make price collusion relatively difficult?* By now, the students recognize that collusion is more difficult when products are differentiated and when information is secret or incomplete. Other factors

complicating collusion include (Church and Ware, pp. 318-325; Waldman and Jensen, pp. 285-291; Lipczynski et al., pp. 169-185):

cost asymmetries
the level of vertical integration
entry barriers
the pace of demand growth
the frequency of sales
elasticity of demand
the pace of innovation
market uncertainty
the social structure of the industry (i.e., the willingness of one firm to act as an organizer)
the strength of antitrust enforcement
the cost of monitoring and detection of cheating.
market concentration (number and size distribution of the firms).

This last factor can be examined quickly by rerunning scenario 4, but with larger groups of competitors (say, ten firms in the market instead of a duopoly). The greater the number of firms, the greater is the likelihood that on will cheat on the cartel to capture the entire market (particularly if the instructor reminds the students that cheating is possible).

TEST OF EFFECTIVENESS

To test the effectiveness of this exercise in teaching the TSB, a Principles of Economics class was split into two groups. The experiment was run near the end of the course, so that the students had the background knowledge required to understand the implications the TSB. The experimental group was taught using the exercise described above, while the control group was taught using traditional teaching methods consisting of a lecture, examples, and Socratic questioning. The time devoted to teaching each group was held constant (40 minutes), and both groups were taught by the same instructor. The next class, five days later, the students understanding of the concepts surrounding the TSB was evaluated by a test consisting of seven multiple choice questions. The results were as follows:

Table 1
TEST OF EFFECTIVENESS RESULTS

	Experimental Group	Control Group
Mean (%)	73.81	66.36
Standard Deviation	19.37	19.89
Observations	42	44

H1 The mean grade of the experimental group is greater than the mean grade of the control group.

The results yield a p-value of 0.041; The null hypothesis is rejected at the 5% significance level. Thus, there is evidence at the 5% level of significance that the mean grade of the experimental group is greater than the mean grade of the control group.

CONCLUSION

The exercise described in this article is an effective method of teaching the Theory of the Second Best and cartel behavior to undergraduate students of economic policy. The participatory nature of the exercise enhances students' understanding of the concepts and provides an association that helps them remember their significance.

REFERENCES

Church, J. & R. Ware (2000). *Industrial organization: A strategic approach*. Toronto: McGraw-Hill.

J M Clark, J.M. (1940) Towards a concept of workable competition. *American Economic Review*, 30(2), 241-256.

Lipczynski, J., J. Wilson & J. Goddard (2009). I*ndustrial organization: Competition, strategy, policy* (*Third Edition*). Toronto: Pearson.

Lipsey, R. & K. Lancaster (1956). The general theory of second best. *The Review of Economic Studies*, 24(1), 11-32.

Ragan, C. (2014). *Economics (Fourteenth Edition)*. Toronto: Pearson.

Waldman, D. & E. Jensen (2013). *Industrial organization: Theory and practice (Fourth Edition*). Montreal: Pearson.

INDUSTRIAL DEVELOPMENT, FINANCIAL DEVELOPMENT AND REGIONAL ECONOMIC GROWTH IN CHINA

Seong-Hoon Lee, Cheongju University
Yonghun Jung, Inha University
Junyeop Lee, Inha University
Sangjoon Lee, Alfred University

ABSTRACT

This study examines the influence of financial development on regional economic growth in China for the period 2003-2010, where industrial development stages are exogenously given and classified as Clark's. Panel analysis of 283 prefectural cities in China for the period shows that financial development is positively associated with regional economic growth in China. Further, the positive association is not found in primary and tertiary sectors while it is found to be strong in secondary or manufacturing sector. Considering China as being in the secondary industrial development stages for the period, the results implies the association of financial development with regional economic growth through the channel of industrial development stages.

INTRODUCTION

This study examines the influence of financial development on regional economic growth in China, where industrial development stages are exogenously given by Clark's classification. China has experienced a rapid economic growth with the financial development. Following the China statistical year book, the real GDP of the Chinese economy grew about 3.1 times from year 2000 to year 2010. Total loans in the financial system grew 3.8 times and total household savings deposited in the financial system grew 4.9 times during the same period. Examining the role of financial development in regional economic growth requires better understanding of the structure of the Chinese economy. Especially manufacturing sector of the Chinese economy has played an important role in the economic growth. While examining China's industry structure, Sasaki and Ueyama (2009) documented that the secondary industry sector composed of about two-thirds of the total output in 2005 and also consisted of almost half of GDP in 2006.

Much attention has been paid in the relationship between economic growth and financial development in recent decades. However, there have been diverse in economic views on the role of financial development in the contribution of economic growth. On the one hand, following Bagehot (1873), Schumpeter (1912), Hicks (1969), Levine (1997) and Miller (1998) argued that financial markets contributed to economic growth. On the other hand, Lucas (1988) shows that the importance of finance in explaining economic growth is "over-stressed."

Levine (1997) suggested five functions of financial systems to accomplish economic growth. These functions are: facilitating risk amelioration, acquiring information about investments and allocating resources, monitoring managers and exerting corporate control, mobilizing savings, and facilitating exchange. These functions facilitate capital accumulation and technological innovation and hence economic growth. Levine (2005) surveyed a large body of empirical studies on the relationship between the financial development and economic growth. A variety of empirical works show that financial development has a positive effect on economic growth. The literature on the finance-growth nexus can be summarized as cross-country, panel, time-series and micro-level analysis.

The cross-country analysis in the early studies shows that financial development enhances economic growth (Goldsmith, 1969; King and Levine, 1993a, 1993b; Levine and Zervos, 1998; La Porta et al., 2002). However, the cross-country regression analysis does not deal with issues of causality and simultaneity bias. Panel analysis mitigates the shortcomings of cross-sectional analysis by taking into account of potential biases from simultaneity, omitted variables and/or country specific effects (Levine, 2005). Still, the results in the panel data studies support a positive effect of financial development on economic growth (Levine, 1998, 1999, 2005; Levine et al., 2000; Beck et al., 2000; Benhabib and Spiegel, 2000; Rousseau and Wachtel, 2000, 2002; Beck and Levine, 2004; Rioja and Valev, 2004a).

Time-series studies that deal with causality issue between finance and growth have in general supported the positive effect of financial development on economic growth (Jung, 1986; King & Levine, 1993a, 1993b; Neusser and Kugler, 1998; Levine et al., 2000; Xu, 2000; Arestis et al., 2001; Khan & Senhadji, 2003; Christopoulos and Tsionas, 2004; Bekaert et al., 2005), even though there are findings of bi-directional causality between finance and growth (Demetriades and Hussein, 1996; Luintel and Khan, 1999; Shan et al., 2001). Micro-level studies with industry level and firm level data, following Rajan and Zingales (1998), also show that financial development is a good indicator to economic growth (Kumar et al., 1999; Wurgler, 2000; Cetorelli and Gambera, 2001; Claessens and Laeven, 2003; Beck et al., 2005).

As an individual country studies, research on the effect of financial development on economic growth in China shows in general its negative association (Boyreau-Debray, 2003; Hasan et al., 2009). However, their analysis uses provincial data for the period before the financial reform in China. Using city-level data focusing on the period after China's financial reforms, especially China's accession to the World Trade Organization WTO in 2001, Zhang et al. (2012) find financial development is positively associated with regional economic growth.

Further decomposing the regional economic growth into Clark's industry classification suggests the role of financial development in facilitating regional economic growth in China be through the channel of industrial development stages. Industries can be classified in diverse ways. A widely known classification of industries into sectors includes primary (extraction of resources such as agriculture, mining, fishing, logging), secondary (production of products from primary industries such as manufacturing, refining, constructing) and tertiary (provision of services) sectors. The patterns of development have some stylized facts in most countries. A country's industrial structure evolves from primary to secondary to tertiary industry sector as its

per capita income grows (Fisher, 1939; Clark, 1940; Kuznets, 1966; Chenery and Syrquin, 1975).

China is in the secondary industrial development stage for the recent decade. As in Table 1, the ratio of secondary industry to GDP in Chana is higher than that in the middle income countries. Thus this paper incorporates China's industry structure into the analysis on the relationship between financial development and the regional economic growth in China. We employ panel data of 283 prefectural cities in China to examine the influence of financial development on regional economic growth and investigate its influence by the Clark's industry classification. Focusing on the period after the recent financial reforms, especially China's entry to World Trade Organization in 2001 and China's establishment of Banking Regulatory Commission in 2003, we use data for the period from 2003 to 2010.

Next section presents empirical analysis and brief conclusions are in section 3.

Table 1
Composition of GDP by Industry (2006)
(%)

	China	Low-income Economies	Middle-income Economies	High-income Economies
Primary Industry	12	20	8	2
Secondary Industry	48	28	37	26
Tertiary Industry	40	52	54	72

Low-income economies are those with a GNI per capita of $905 or less and high-income economies are those with a GNI per capita of $11,116 or more in 2006. Composition of GDP by industry for each income bracket is calculated as a weighted average of the corresponding economies. China's GNI per capita in 2006 is $2,000.
Source: Sasaki and Ueyama (2009); World Bank *"World Development Indicators."*

EMPIRICAL ANALYSIS

As in Zhang et al. (2012), we employ a panel of Chinese city data focusing on the period after the recent financial reforms including China's entry to World Trade Organization in 2001 and China's establishment of Banking Regulatory Commission in 2003. The Chinese government classifies cities as three categories such as municipalities, prefecture cities and county cities. The central government directly governs four municipalities (Beijing, Shanghai, Tianjin and Chongqing) and the provincial government directly governs each province that consists of about 10 prefecture cities which governs county cities. To analyze the effects of financial development on the regional economic growth in China, a panel of 286 Chinese prefecture cities is used from the China City Statistical Yearbook for the period 2003-2010. Table 2 shows the descriptive statistics for the data used in this study. On the average, growth rate, which is per capita real GDP growth rate measured by log difference is highest in primary sector and then in secondary sector.

Table 2
Descriptive statistics

Variable	N	Average	Std
Growth rate			
All sectors	1969	0.129	0.091
Primary sector	1955	0.178	0.441
Secondary sector	1967	0.133	0.168
Tertiary sector	1967	0.120	0.139
FD	2254	1.142	0.507
STE	2256	0.032	0.065
GOV	2254	0.132	0.089
FDI	2168	0.0003	0.0004
SOC	2253	0.005	0.002
Road	2246	0.151	0.362

Growth rate is per capita real GDP growth rate measured by log difference, FD is a financial depth measured by the ratio of total household savings deposited in the financial system to GDP (savings/GDP), STE is a science technology expenditure over government expenditure, FDI is the ratio of real foreign direct investment over real GDP, GOV is the share of government expenditure over GDP, SOC is the ratio of real social overhead capital investment over real GDP, Road is the density of road measured by the total area of road over the total area of land.

The model employed in this study is an extension of the models in King and Levine (1993) and Zhang et al. (2012). The effects of financial development on the economic growth are estimated by the Clark's industry classification. We include control variables such as year and city dummies, science technology expenditure over government expenditure (STE), the ratio of real foreign direct investment over real GDP (FDI), the share of government expenditure over GDP (GOV), the ratio of real social overhead capital investment over real GDP (SOC), the density of road measured by the total area of road over the total area of land (Road). All the independent variables are used by one year lagged log values to avoid simultaneity bias problems. Thus, our estimation is based on the following econometric model:.

$$y_{ijt} = \beta_{0j} + \beta_{1j} FD_{it-1} + \gamma_j X_{it-1} + \mu_{ij} + \varepsilon_{tj} + \tau_{ijt},$$

where y is the per capita real GDP growth rate in city i at time t for each j sector, FD is a financial depth measured by the ratio of total household savings deposited in the financial system to GDP. Industry sectors are classified as primary, secondary and tertiary sectors and control variables of X vector include STE, FDI, GOV, SOC and Road variables. Finally, μ and ε indicate city-specific and year-specific errors, respectively, and τ is the remaining identically independent error terms.

Table 3
Results with two-way fixed effects

	Per capital real GDP growth			
	All sectors	Primary sector	Secondary sector	Tertiary sector
FD	0.0770***	-0.0173	0.142***	0.0418
	(0.0180)	(0.112)	(0.0391)	(0.0270)
STE	0.00386	-0.00804	-0.00472	0.00509
	(0.00364)	(0.0226)	(0.00785)	(0.00542)
FDI	-0.00520**	0.0154	-0.0128**	-0.00918***
	(0.00230)	(0.0142)	(0.00496)	(0.00342)
GOV	0.0939***	0.0761	0.149***	0.0560**
	(0.0182)	(0.113)	(0.0393)	(0.0271)
SOC	0.0633***	0.00201	0.0441*	0.0567***
	(0.0106)	(0.0660)	(0.0230)	(0.0159)
Road	0.00903	-0.0328	0.0148	-0.00038
	(0.00692)	(0.0429)	(0.0149)	(0.0103)
N	1873	1859	1871	1871
R^2	0.314	0.167	0.355	0.340

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
The estimations are based on two-way fixed effects including time and city dummy variables. FD is a financial depth measured by the ratio of total household savings deposited in the financial system to GDP (savings/GDP), STE is a science technology expenditure over government expenditure, FDI is the ratio of real foreign direct investment over real GDP, GOV is the share of government expenditure over GDP, SOC is the ratio of real social overhead capital investment over real GDP, Road is the density of road measured by the total area of road over the total area of land. All the independent variables except year dummies are used by one year lagged log values.

Table 3 shows empirical results of two-way fixed effect panel estimations where city-specific and time-specific dummy variables are included. The results indicate that financial development in all sectors has a significantly positive impact on regional economic growth. However, estimation by the Clark's industry classification shows interesting results. The effects of financial development on the regional economic growth are not statistically significant in both the primary and the tertiary sectors. The significantly positive role of financial development in explaining the regional economic growth is shown only in the secondary sector at the 5% significance level. Thus financial development is a strong contributor to the regional economic growth in secondary, especially manufacturing sector for the recent decade in China. Considering China as being in the secondary industrial development stages for the period, the results implies the association of financial development with regional economic growth through the channel of industrial development stages. Other contributors to the regional economic growth in China are the share of government expenditure over GDP and the ratio of real social overhead capital investment over real GDP. That is, the larger the government expenditure and the more the real social overhead capital investment, the higher will be the regional economic growth.

CONCLUSION

We examined the effect of financial development on regional economic growth in China for the period 2003-2010, where industrial development stages are exogenously given in a way of Clark's classification. Extending the work by Zhang et al. (2012), we employed a panel of Chinese city data focusing on the period after the recent financial reforms of China's establishment of Banking Regulatory Commission in 2003. Two-way fixed effect panel analysis of 283 prefectural cities in China for the period showed that financial development is positively associated with regional economic growth in China. Further, the positive association was not statistically significant in primary and tertiary sectors while it was strong significant and positive in secondary or manufacturing sector.

The results indicate the positive contribution of financial development to regional economic growth has been played through the channel of industrial development stages, especially the secondary sector in China for the period 2003-2010. Thus the results suggest economic and financial policies in China, of fostering financial development, should be in conjunction with the manufacturing industrial sector, rather than primary and tertiary sectors, to enhance the regional economic growth. Our findings also imply that a theoretical dynamic model of a dual economy, consisting of agricultural and manufacturing sectors in association with financial development, is plausible in Chinese economy. In addition, we found that the government expenditure and the real social overhead capital investment were positively associated with the regional economic growth in China for recent decade.

REFERENCES

Arestis, Philip, Demetriades, Panicos O., Luintel, Kul B., 2001. Financial development and economic growth: the role of stock markets. Journal of Money, Credit and Banking 33 (1), 16–41.

Bagehot, Walter. 1873. Lombard street. Homewood, IL: Richard D. Irwin, 1962 Edition.

Beck, Thorsten, Levine, Ross, 2004. Stock markets, banks, and growth: panel evidence. Journal of Banking & Finance 28 (3), 423–442.

Beck, Thorsten, Levine, Ross, Loayza, Norman, 2000. Finance and the source of growth. Journal of Financial Economics 58 (1–2), 261–300.

Beck, Thorsten, Demirgüç-Kunt, Asli, Maksimovic, Vojislav, 2005. Financial and legal constraints to growth: does firm size matter? Journal of Finance 60 (1), 137–177.

Bekaert, Geert, Harvey, Campbell R., Lundblad, Christian, 2005. Does financial liberalization spur growth? Journal of Financial Economics 77 (1), 3–55.

Benhabib, Jess, Spiegel, Mark M., 2000. The role of financial development in growth and investment. Journal of Economic Growth 5 (4), 341–360.

Boyreau-Debray, Genevieve, 2003. Financial Intermediation and Growth: Chinese Style. The World Bank, Policy Research Working Paper Series 3027.

Carkovic, Maria, Levine, Ross, 2005. Does Foreign Direct Investment Accelerate Economic Growth? Does Foreign Direct Investment Promote Development. Institute for International Economics, Washington, DC, US, pp. 195–220.

Cetorelli, Nicola, Gambera, Michele, 2001. Banking market structure, financial dependence and growth: international evidence from industry data. Journal of Finance 56 (2), 617–648.

Chenery, H. B., and M. Syrquin. 1975. Patterns of Development, 1957–1970. Oxford University Press, London.

Christopoulos, Dimitris K., Tsionas, Efthymios G., 2004. Financial development and economic growth: evidence from panel unit root and cointegration tests. Journal of Development Economics 73 (1), 55–74.

Clark, C. 1940. The Conditions of Economic Progress. McMillan & Co.s London.

Demetriades, Panicos O., Hussein, Khaled, A., 1996. Does financial development cause economic growth? Time-series evidence from 16 countries. Journal of Development Economics 51 (2), 387–411.

Fisher, A.G.B.1939. "Primary, Secondary and Tertiary Production," Economic Record 15, June, 24–38.

Goldsmith, Raymond W., 1969. Financial Structure and Development. Yale University Press, New Haven, US.

Hasan, Iftekhar, Wachtel, Paul, Zhou, Mingming, 2009. Institutional development, financial deepening and economic growth: evidence from China. Journal of Banking & Finance 33 (1), 157–170.

Hicks, John. 1969. A theory of economic history. Oxford: Clarendon Press.

Jung, Woo S., 1986. Financial development and economic growth: international evidence. Economic Development and Cultural Change 34 (2), 333–346.

Khan, M., Senhadji, A. 2003. Financial development and economic growth: A review and new evidence. Journal of African Economies, 12, ii89–ii110.

King, Robert G., Levine, Ross, 1993. Finance and growth: Schumpeter might be right. Quarterly Journal of Economics 108 (3), 717–738.

Kumar, Krishna, B., Rajan, Raghuram, G., Zingales, Luigi, 1999. What Determines Firm Size? National Bureau of Economic Research, Working Paper No. 7208.

Kuznets, S. 1966. Modern Economic Growth, Yale University Press, New Haven, CT.

La Porta, Rafael, Lopez-de-Silanes, Florencio, Shleifer, Andrei, 2002. Government ownership of banks. Journal of Finance 57 (1), 265–301.

Levine, Ross, 1998. The legal environment, banks, and long-run economic growth. Journal of Money, Credit and Banking 30 (3), 596–613.

Levine, Ross, 1999. Law, finance, and economic growth. Journal of Financial Intermediation 8 (1–2), 8–35.

Levine, Ross. 2005. Finance and Growth: Theory and Evidence. Handbook of Economic Growth, vol. 1, no. 1. Elsevier, Amsterdam, Netherlands, pp. 865–934 (Chapter 12).

Levine, Ross, Zervos, Sara, 1998. Stock markets, banks, and economic growth. American Economic Review 88 (3), 537–558.

Levine, Ross, Loayza, Norman, Beck, Thorsten, 2000. Financial intermediation and growth: causality and causes. Journal of Monetary Economics 46 (1), 31–77.

Lucas Jr., Robert E., 1988. On the mechanics of economic development. Journal of Monetary Economics 22, 3–42.

Luintel, B. K., Khan, M. 1999. A quantitative re-assessment of the finance-growth nexus: Evidence from amultivariate VAR. Journal of Development Economics, 60, 381–405.

Miller, Merton H., 1998. Financial markets and economic growth. Journal of Applied Corporate Finance 11 (3), 8–15.

Neusser, Klaus, Kugler, Maurice, 1998. Manufacturing growth and financial development: evidence from OECD countries. Review of Economics and Statistics 80 (4), 638–646.

Rajan, Raghuram G., Zingales, Luigi, 1998. Financial dependence and growth. American Economic Review 88 (3), 559–586.

Rioja, Felix, Valev, Neven, 2004. Finance and the sources of growth at various stages of economic development. Economic Inquiry 42 (1), 127–140.

Rousseau, Peter L., Wachtel, Paul, 2000. Equity markets and growth: cross-country evidence on timing and outcomes, 1980–1995. Journal of Banking & Finance 24 (12), 1933–1957.

Rousseau, Peter L., Wachtel, Paul, 2002. Inflation thresholds and the finance-growth nexus. Journal of International Money and Finance 21 (6), 777–793.

Sasaki, Hitoshi, Ueyama, Satoko. 2009. "China's Industrial Structure and its Changes in Recent Years: An Analysis of the 1997–2005 Input-Output Tables," Bank of Japan No.09-E-2, June.

Schumpeter, J. 1912. The Theory of Economic Development. Cambridge, MA: Harvard Univ.

Shan, J., Morris, A., Sun, F. 2001. Financial development and economic growth: An egg-chicken problem? Review of International Economics, 9, 443–454.

Wurgler, Jeffrey, 2000. Financial markets and the allocation of capital. Journal of Financial Economics 58 (1–2), 187–214.

Xu, Zhenhui, 2000. Financial development, investment, and economic growth. Economic Inquiry 38 (2), 331–344.

Zhang, Jin, Wang, Lanfang, Wang, Susheng. 2012. Financial development and economic growth: Recent evidence from China. Journal of Comparative Economics 40, 393–412.

A REPEATED CROSS-SECTION AND PSEUDO PANEL ANALYSIS OF ALLEVIATING POVERTY IN DEVELOPING ECONOMIES: THE PHILIPPINE CASE[1]

John Paolo R. Rivera[2], Asian Institute of Management

ABSTRACT

Government-sponsored programs may have overlooked the root causes of poverty. The concepts of transient and chronic poverty demonstrate the behavior of households towards their socio-economic status. We used a repeated cross-section and pseudo panel analysis to estimate the probability that a household will move from the state of being poor to non-poor, non-poor to poor, or stay in the status-quo. Results showed that there are households who have the tendency to remain on their current state while some moved from one state to another.

Keywords: chronic poverty, pseudo panel, repeated cross section analysis, transient poverty

INTRODUCTION

Poverty has been unrelenting in all economies. It is the basket case of most economies particularly among Southeast Asian nations. Poverty alleviation has been a principal goal for most economies because it has incessantly posed a long-term struggle, especially to the Philippines, since this has been the primary objective of all administrations. As per Schelzig (2005), the Philippines has a perceptible unequal income distribution which supports the premise that Filipinos in the lower distribution are highly susceptible to impoverished living conditions leaving households vulnerable. Meanwhile, according to Albert and Ramos (2010), income shocks have debilitating effects especially to the poor, which drives households to engage themselves in risky strategies that sometimes have negative effects that are irreversible and eventually succumb to deeper poverty.

The International Labor Organization (ILO), as cited by Schelzig (2005), listed five non-monetary categories that define whether people are poor – food, water and sanitation, health, education and shelter. In 2009, according to the National Statistics Office (NSO), the share of food to total family expenditures accounts for 42.6 percent, which is a considerable portion of income allocation and signifies as one of the priorities of consumption spending. With the existing income inequality, this can be translated to food inequality through the income channel, which means that families and individuals in the lower income distribution are unable to gain access to food because of the lack of ability to afford decent food consumption. This is a serious issue because according to Reyes (2001), the poorest Filipino households allocate a significant portion of their income on food. Decreasing real income also signifies that capacity to spend on

food is restricted; hence, families are forced to concentrate household expenditure on basic necessities.

The National Statistical Coordination Board (NSCB) and NSO come up with measures to assess the depth of poverty in the country. Measures such as poverty incidence, Gini coefficient, and income and expenditure ratios all relate to the traditional measure of welfare, which is the level of income. The measure of welfare and poverty is not limited to income and expenditure alone. In the Philippines, there are two official measures of poverty, namely, the food threshold and the poverty threshold (Schelzig, 2005). Moreover, Pedro, Candelaria, Velasco and Barba (n.d.) estimated food threshold adjusted to the lower 30 percent of the income distribution to represent the poor in the population to gauge poverty incidence through food.

However, beyond economic factors, there are other explanations as to why poverty exists in the Philippines. According to Abad and Eviota (1983), poverty is a condition caused by exhibiting anti-development traits, values and attitudes such as refusal for improvement and resistance to change, which implies that the poverty-stricken cause their own predicament for they are responsible for their own behaviors. Further, in a study by Spears (2010), the poor developed a set of beliefs and behaviors that are adaptive but constraining to poverty, which then creates a widened culture of poverty and as stated by Abad and Eviota (1983), such a culture will continue to perpetuate in the succeeding years. In addition, according to Bennett (2008), it is highly debated upon why the poor people continue to behave irrational ways that limits them in their impoverished state. Ultimately, in order to understand the Philippine poverty situation and why it is non-improving, the culture of the country's poverty can be evaluated by the Filipino characteristics inherent to the majority. It is then of uttermost importance to determine whether the Filipino culture may then be a pro-poverty culture.

Hence, we investigate the movement of household in and out of the poverty threshold. Determining this will allow us to provide an illustration on why households become impoverish, or continuously remain under poverty, or eventually escape poverty and remain non-poor for an extended period of time. This can be determined by knowing how households switch their spending behavior if their income and/or wealth change. Hence, our main research problem: "Is there any significant probability that a poor household can escape poverty despite the inadequacies of government sponsored programs to alleviate poverty?" To address these problems, the following objectives are set:

- *To identify the probability that a household will stay on their current socio-economic status or move to another state;*
- *To generate recommendations on how poor households can increase the probability of moving out of poverty.*

Accomplishing these objectives will direct us to the root cause of poverty, which is the lopsided spending of households with respect to the income that they generate. Likewise, despite the provision of government-sponsored programs to address poverty, we will be able to argue that it is necessary but not sufficient to address poverty because the role of households in managing their finances is also necessary. This study is imperative to the government in their

formulation of poverty-alleviation measures because a significant amount of funds are allocated to these programs, which are still deemed to be inadequate by its recipients.

This study is also essential to households because this will provide a perspective on how should they manage their finances in order to motivate a rational spending behavior. Upon evaluation, results of this study can provide a framework to policymakers to address program implementation more efficiently and to suggest programs that can actually sustain household welfare and answer poverty issues.

POVERTY IN THE PHILIPPINES

With the recent global financial crisis, continuous severe natural calamities, and rising fuel and food prices, the government's goal of reducing poverty is becoming more difficult since these situations have been pulling more people into poverty. The official poverty statistics from the NSCB, as seen from the Table 1, reveal that the annual per capita poverty threshold have increased by 26.2 percent from 2006 to 2009. Although poverty incidence among households declined from 2006 to 2009, the poverty incidence among the population has increased by 0.1 percentage point and is still relatively high compared to that of Indonesia, Malaysia, Thailand, and Viet Nam as per the reports from the United Nations Development Program (UNDP). With the economy's rising population growth is the growing number of the poor. The magnitude of poor people has been increasing from 19.8 million people in 2003 to 23.14 million people in 2009, an additional 3.34 million poor people in six years. There is also an additional 570,000 poor families, from 3.29 million poor families in 2003 to 3.86 million poor families in 2009.

The other three poverty measures have decreased from 2006 to 2009 implying a better condition. The income gap[3] and the poverty gap[4] have been reduced by 1.5 points and 3 points in three years, respectively. The severity of poverty or the total of the squared income shortfall has also declined by 0.2 points in 2009.

One characteristic of the economy's poverty, as pointed out by Reyes, Tabuga, Mina, Asis and Datu (2010) is the inequality across regions. Figure 1 shows the thematic maps of 2009 income gap, poverty gap, and severity of poverty where red shades show higher gaps and are therefore comparatively worse off areas that green shaded areas. The darker the red, the situation in that area is worse off compared to the rest and the darker the green shade, the better off.

There are significantly higher poverty measures in Eastern Visayas and Caraga Administrative Region where Eastern Samar and Agusan del Sur are located, respectively, compared to that of the National Capital Region (NCR). Caraga is one of the most impoverished regions whose primary source of income is the agriculture and forestry sector similar to the Eastern Visayas. Zamboanga Peninsula is also another region stricken with more poverty compared with the Autonomous Region of Muslim Mindanao (ARMM).

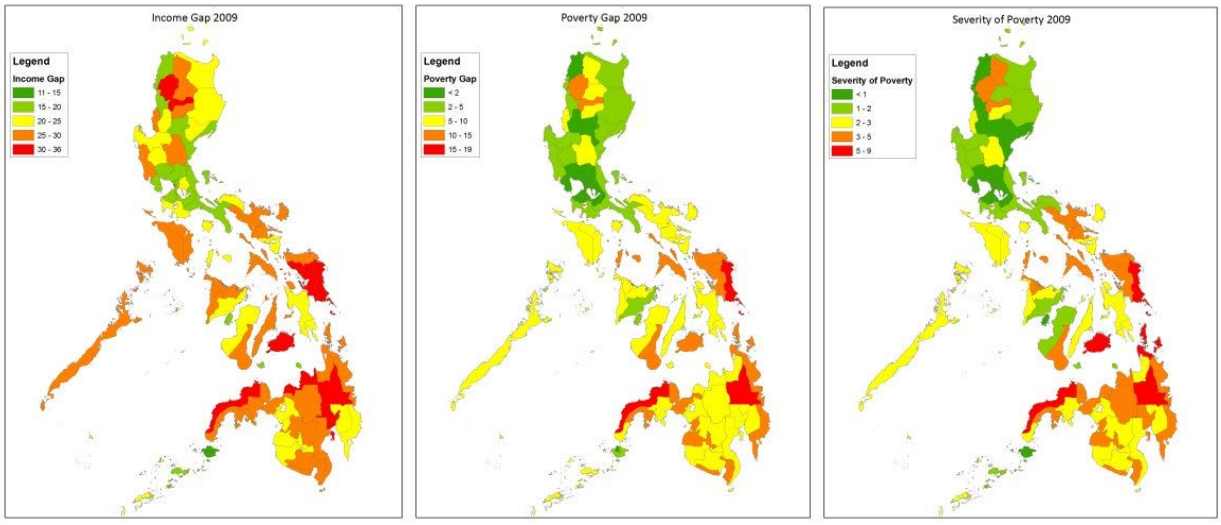
Table 1
ANNUAL PER CAPITA THRESHOLDS AND POVERTY INCIDENCE IN THE PHILIPPINES

Year	2003	2006	2009
Annual Per Capita Poverty Threshold (PHP)	10,976	13,348	16,841
Poverty Incidence (%)			
Families	20	21.1	20.9
Population	24.9	26.4	26.5
Magnitude of poor (in million)			
Families	3.29	3.67	3.86
Population	19.8	22.17	23.14
Subsistence Incidence (%)			
Families	8.2	8.7	7.9
Population	11.1	11.7	10.8
Magnitude of subsistence poor (in million)			
Families	1.36	1.51	1.45
Population	8.8	9.85	9.44

Source: National Statistical Coordination Board (NSCB)

The map emphasizes the reality that poverty in the Philippines is a geographical matter that calls for more programs in alleviating poverty prioritizing regions with significantly worse conditions. As reported by the 2009 Philippine Poverty Statistics, the poverty incidence among families in NCR has improved from 3.4 in 2006 to 2.6 in 2009 but the considerably different poverty incidence among families in the ARMM has worsened from 36.5 in 2006 to 38.1 in 2009 and has been the region with most poverty incidence in 2006 and 2009.

Figure 1
THEMATIC MAP OF 2009 INCOME GAP, POVERTY GAP, SEVERITY OF POVERTY



Source: National Statistical Coordination Board (NSCB)

Table 2 INCOME GAP, POVERTY GAP AND SEVERITY OF POVERTY			
Year	2003	2006	2009
Income Gap	27.7	27.2	25.7
Poverty Gap	5.6	5.7	2.7
Severity of Poverty	2.2	2.2	2
Source: National Statistical Coordination Board (NSCB)			

National averages do not show the staggering urban and rural differences and also the regional variations. Regional averages also do not show provincial disparities (Schelzig, 2005). The startling provincial level differences are illustrated in Table 3. Eastern Samar and Agusan del Sur's, poverty incidence among population is 54.0 and 58.1, respectively, both more than 20 times higher than that of NCR District IV's poverty incidence.

Attention should also be paid to the fact that both provinces have worsened since 2006 in both poverty incidence among families and population when NCR District IV have significantly improved. The complete details of the measures of poverty incidence for the Philippines and its regions are shown in Table 2, Table 3, and Table 4.

This just confirms that there is certainly a need to employ improved and well-designed policies that takes into account provincial and regional profile for a more strategic distribution to potential key areas for a more socially and economically equal society.

The Sources of Poverty

Developing countries are characterized by having a high income inequality, inequitable distribution of income, and poverty incidence (Todaro & Smith, 2008). Although the Philippines has been vigilant in addressing poverty, it has been sluggish compared to other Asian economies that have been successful in reducing their country's poverty incidence. Even Cambodia, Indonesia, Malaysia, and Vietnam whose annual real Gross Domestic Product (GDP) growth rate is lower than the Philippines, as per the reports of the UNDP, have outperformed the country in reducing poverty for the past two decades. The general literature on poverty in the Philippines cited that the main reason for the low reduction of poverty is the slow growth of the economy. With this reduced economic expansion accompanied by high population growth rate, the country suffers a slow growth rate of per capita income. However, even though the economy will experience high growth, the quality of this growth is vital since not all components of economic growth is in support of the poor.

Table 3
POVERTY INCIDENCE IN SELECTED PROVINCES
(NCR, EASTERN SAMAR, AND AGUSAN DEL SUR)

Year / Province	2003	2006	2009	2003	2006	2009	2003	2006	2009
	Annual Per Capita Poverty Threshold			Poverty Incidence Among Families (%)			Poverty Incidence Among Population (%)		
NCR District IV	13,997	16,487	19,802	1.8	2.9	1.6	2.7	5.0	2.5
Eastern Samar	10,106	12,195	16,385	29.8	37.6	45.8	36.4	47.8	54.0
Agusan Del Sur	11,226	14,004	18,443	48.5	45.5	51.2	56.0	53.9	58.1
				Magnitude of Poor Families			Magnitude of Poor Population		
NCR District				10,769	17,942	12,389	78,834	145,819	88,850
Eastern Samar				22,642	31,165	41,359	141,236	206,979	237,122
Agusan Del Sur				54,915	54,433	57,189	313,709	319,936	343,060

Source: National Statistical Coordination Board (NSCB)

Furthermore, according to Aldaba (2009), since the economy fails to maintain a high level of sustained growth, it cannot create the necessary employment, therefore, inadequate income for the poor that further reduces their opportunity to fight poverty thus increasing the gap between the rich and the poor.

Another contributory factor of insistent poverty is the high level of population growth. High level of population growth may indicate increasing family size and the larger the family size is, the greater the household or family's probability of being poor (Schelzig, 2009) because, as assessed by Orbeta (2002), high fertility is related to the decline in human capital investments. An additional member of a family means an additional allocation of a usually meager income, thus lesser division of family resources (Schelzig, 2005), most particularly, food and nutrition.

Moreover, the rapid population growth impedes economic development for two interconnected reasons. First, rapid population growth lessens per capita income, since the people, especially the poor cannot sacrifice basic commodities, their savings are reduced and so will be the resource for investment in productive capacity. This will sequentially decrease overall economic growth and increase poverty (Schelzig, 2005). Second, the country's large population that is continuously growing exceeds the capacity of the industry to absorb new labor. Since, there is a rapid increase in the labor force, and again, with the lack of quality employment, the outcome will be more unemployed individuals negatively affecting the development of the economy. This insufficient generation of employment and low quality of employment is another reason for the persistent high poverty incidence. The availability of employment cannot keep up with the growth pace of the labor force. In a matter of a decade, the country's labor force has increased by more than 50 percent and even the total labor force participation increased due to the higher participation of women in the labor force. Even with Filipinos opting to work overseas, according to Aldaba (2009), unemployment rates are still high.

Increasing population is not an issue and will not contribute to poverty if all required resources are able to deal with the additional population such as employment and the country's funds. However, the Philippines has many poor households; and with persistent government budget deficits and increasing labor force, the rapid population growth is now a problem that must be addressed (Schelzig, 2005).

In 2010, according to NSO, the share of employment in agriculture to the total employment is 33 percent and most of the profiles of employees working in this sector including the industry sectors are considered poor. The Annual Poverty Indicator Survey (APIS) of the NSO, using the bottom 40 percent income range as a proxy for the poor, revealed that more than half of the poor are employed in agriculture in 2009 and are usually laborers and farmers (Schelzig, 2005). However, the agricultural sector that has been overlooked and has not been given proper management is lacking sustainable and quality employment (Aldaba, 2009). As argued by Schelzig (2005), the link between poverty and employment is principally apparent, not in the unemployment, but in the quality of the employment since the poor are generally working in jobs with low income and low productivity. Thus, even though the poor are employed, they are not working in a quality environment. Furthermore, rural areas where most of the poor are rely on the agricultural sector. Very little were done to strengthen, develop, and reduce or eradicate market distortions of this sector.

Table 4
ANNUAL PER CAPITA POVERTY THRESHOLD AND POVERTY INCIDENCE AMONG FAMILIES

Region	Annual Per Capita Poverty Threshold (PHP)			Poverty Incidence among Families						Share to Total Poor Families		
				Estimate			Coefficient of Variation					
	2003	2006	2009	2003	2006	2009	2003	2006	2009	2003	2006	2009
Philippines	10,976	13,348	16,841	20	21.1	20.9	2.3	2.3	2.1	100	100	100
NCR	13,997	16,487	19,802	2.1	3.4	2.6	12	13.1	12	1.5	2.2	1.7
CAR	10,881	12,976	16,122	16.1	18.6	17.1	11.2	12.8	11.8	1.4	1.5	1.4
Ilocos	11,791	14,350	17,768	17.8	20.4	17.8	7.8	7.6	7	4.7	5.3	4.6
Cagayan Valley	10,350	12,212	15,306	15.2	15.5	14.5	8.4	9.2	8.3	2.7	2.6	2.4
Central Luzon	12,771	15,374	18,981	9.4	12	12	8.6	8.4	7.6	5.2	6.2	6.3
CALABAR ZON	12,394	14,284	17,779	9.2	9.4	10.3	8.7	10.6	7.9	6.1	5.7	6.4
MIMA ROPA	10,398	12,610	15,769	29.8	34.3	27.6	6.6	7.2	6.8	4.5	5.1	4.2
Bicol	11,476	13,645	17,146	38	36.1	36	4.9	5	4.2	10.9	9.9	10
Western Visayas	10,548	12,432	16,036	23.5	22.1	23.8	6.7	7.2	6.4	9.1	8.2	9
Central Visayas	11,798	14,468	17,848	32.1	33.5	30.2	6.5	6.4	6.3	11.8	11.8	10.8
Eastern Visayas	9,850	11,885	15,910	30.2	31.1	33.2	5.9	5.8	5.3	6.9	6.9	7.4
Zamboanga Peninsula	9,642	11,810	15,160	40.5	34.2	36.6	7.1	9.3	6.7	7.2	6.1	6.3

Table 4
ANNUAL PER CAPITA POVERTY THRESHOLD AND POVERTY INCIDENCE AMONG FAMILIES

Region	Annual Per Capita Poverty Threshold (PHP)			Poverty Incidence among Families						Share to Total Poor Families		
				Estimate			Coefficient of Variation					
	2003	2006	2009	2003	2006	2009	2003	2006	2009	2003	2006	2009
Northern Mindanao	10,501	12,987	16,568	32.4	32.7	32.8	7.2	5.7	5.8	7.3	7	7.1
Davao	10,737	13,469	17,040	25.4	26.2	25.6	8.3	8.7	8.4	6.3	6	5.9
SOCCSK SARGEN	10,277	12,530	15,762	27.2	27.1	28.1	8.1	7	6.7	5.8	5.5	5.8
Caraga	10,355	12,935	16,858	37.6	36.9	39.8	6.2	6.4	5.4	4.7	4.5	4.9
ARMM	9,664	12,358	16,334	25	36.5	38.1	10.1	7.4	6.1	3.8	5.3	5.7

Source: National Statistical Coordination Board (NSCB)

According to Aldaba (2009), there has been less interest in advancing agricultural productivity and less concern in developing the necessary infrastructure to increase productivity further decreasing the income opportunities especially for poor in the rural areas. If these sectors are improved, it will create more quality jobs to individuals who need it most.

A financial crisis may also be a source of poverty, but in the case of the Philippines, it did not generate poverty but has contributed to the continuously slow pace of poverty reduction. The 1997 Asian Financial Crisis (AFC) became a hindrance to poverty alleviation and with another global crisis is another impediment to poverty reduction. With this 2008 Global Financial Crisis (GFC), the expansion of the Philippine economy slowed but the impact to the total economy was not greatly detrimental compared to the impact to other Southeast Asian economies. The economy's banking system was better prepared for the current crisis than they were on the Asian financial crisis. However, the current crisis has greatly affected some sectors of the economy than it has affected other sectors such as the trade and manufacturing sectors. Furthermore, the high inflation rates brought about by the crisis have also lessened the positive effect of economic growth on poverty (Aldaba, 2009).

Using an alternative price index, in addition, the poor deals with a higher inflation rate than the country's official rate driving even more individuals to poverty when the price of food increases. It also drives more poor people to worse living standards since they have to allocate more of their income in consuming food, thus, sacrificing education and health care (Son, 2008).

Another contributory factor is the high and persistent levels of economic inequality such as income, welfare and asset inequality that diminishes the positive effect of economic growth. A study by Deininger and Squire (1998) argued that an economy's initial land distribution has an effect on the succeeding expansion of its economy and performance of human development. If an economy has a high land inequality initially, it is likely that it will exhibit a lower income growth rate in the long-term and slower rate of alleviating poverty than an economy with more equitable land distribution initially. Land inequitable distribution has been the country's problem for over four decades.

There is also a high-income inequality in the Philippines. In 2006, NSCB reported that the income share of the bottom 10 percent of the economy's population is at 1.86. The Gini

coefficient in 2006 is 45.8. Based on the statistics from the UNDP, although the unequal distribution has improved from the 46.8 in 1991 and the extent of the unequal distribution of income is better than that of Malaysia (46.2 in 2009), Thailand (53.6 in 2009) and Singapore (47.8 in 2009), it is still very high. Furthermore, Cambodia (est. 44.4 in 2007), Indonesia (36.8 in 2009), Laos (36.7 in 2008) and Viet Nam (37.6 in 2008) fared better in income distribution.

One problem that high-income inequality also conveys, as suggested by Schelzig (2005), is that measures for poverty are extremely sensitive to the poverty threshold because of the large number of individuals in the initial deciles. A slight adjustment to the poverty line can have considerable adjustments to number of individuals considered as poor (Schelzig, 2005).

As emphasized, poverty in the country is a geographical matter since there is a wide disparity in the standards of living and human development in the different regions. This interregional and intra-regional inequality contributes to poverty. According to Balisacan (2003) and Aldaba (2009), intra-regional inequality contributes 82 percent of overall inequality; thus, poverty, implying that state policies must improve distribution within the regions.

Other causes of poverty asserted by Aldaba (2009) are the recurrent shocks to the country and exposure to risks such as financial crises, natural disasters, social conflicts, and environmental property. Social conflicts worsen poverty incidence since these hinder individuals from engaging in their economic activities as they are removed from their homes and their sources of income. These do not just result to displacement but to disablement and deaths, therefore, loss of household heads and increase in dependency ratios. Social conflicts also disrupt people's access to basic services and devastate transport systems and life in the rural areas. In brief, it affects individual's access to all forms of capital – physical, natural, social, financial, and human (Schelzig, 2005). Natural disaster and environmental property also result to higher poverty incidences because occurrence of devastating calamities affects mostly standard of living of the poor. They experience more losses since their sources of income and health is most likely reliant on the environment (Aldaba, 2009).

Aside from the abovementioned economic factors that cause poverty in the Philippines, it is also imperative to look at behavioral and social factors. In an empowerment study by Irarrazaval (1995), personal initiative and responsible work were also relevant factors for an individual's economic success as well as the poverty causing factors of an individual with the highest frequencies include laziness and lack of initiative. Hine, Montiel, Cooksey and Lewko (2005) further affirmed this in their study that considered laziness and the lack of effort as a causal factor of poverty. According to Kim (2011), the Spaniards mocked the Filipinos as "*perezoso indios"* meaning "lazy Filipinos" in which these insults perpetuate and are generally true up to day. In relation, Andres and Ilada-Andres (1998) stated that Filipinos are known to be "*Juan Tamad*" or lazy *Juan* to illustrate the tendency of Filipinos to put off work or even put off looking for work that is counterproductive to success due to the time and opportunities wasted. Moreover, according to Ilda (2011), this laziness is attributed to the lack of drive and the reliance of the poor people on the other sectors to uplift their economic situation.

On the other hand, Ilda (2011) stated that instant gratification, defined as the happiness resulting from impulsive decisions, is one of the traits the most Filipinos are addicted to which explains why most Filipinos habitually execute any task the quickest and easiest way out.

Additionally, this instant gratification is supported by the collective notion that most Filipinos are also impatient. That is, according to Irving Fisher, as cited by Spears (2010), a lower level of income will likely entail a higher tendency of impatience, other things held constant. Moreover, as identified by Ilda (2011), this behavior also implies the collective need for Filipinos to blame others for the lack of progress and development they are experiencing in their personal lives. It follows that the Filipinos, mostly the poor, blame the government for the lack of job opportunities and lifestyle conditions and are dead set on waiting for these so called projects. Similarly, in article penned by Bennett (2008), it states that as the government aids the poor more, the latter are less likely to work for themselves to provide their own necessities. Furthermore, according to Ilda (2011), those who live in the slum areas have a dependent mentality wherein they strongly believe they are entitled to receive government benefits and privileges instead of being responsible for their own lives. Another example of which is the keenness of Filipinos for gambling. Consequentially, their belief in instant gratification and heavy dependence cause them to not believe in hard work and patience, and their lack of contribution to the betterment of their own happiness makes them useless members of society.

Chronic and Transitory Poverty

Even if an economy is considered developed there still exists a significant amount of individuals who succumbs into low socio-economic status. Analyzing its dynamics, although constrained due to data availability, will provide significant understanding to why households remain or move in and out of poverty. Households who have been persistently in a state of being poor over a long period of time are in chronic poverty while those in transient poverty are those who are poor only for a time or those who shift in and out of poverty. Barrientos, et al. (2005), as cited in Ribas and Machado (2007) categorized the definitions of chronic and transient poverty in the literature into three groups. The first category, which this paper utilizes, focuses on the duration of being in a state of poverty. It is chronic poor when the levels of their per capita income or consumption are constantly below the poverty line while it is considered transient when the levels fluctuate at the poverty line (Gaiha & Deolalikar, 1993). The second category focuses on the components of income or consumption. The constant component is the determinant of chronic poverty while the fluctuating component is that of transient poverty (Jalan & Ravallion, 1998). The third category focuses on current income and its variability among households to assess vulnerability to poverty (Pritchett, et al., 2000).

One importance of identifying poverty between chronic and transient is verifying the differences in the determinants or components commonly associated between the two including their behavior on income generation and spending. Some components may be more associated to chronic poverty than transient poverty, thus may also present disparity in addressing them. This will, therefore, lead to government programs being weakly implemented because these are overlooking the appropriate target. Some of these programs may address transient poverty while miss tackling chronic poverty suggesting why interventions are not effective.

The characteristics most commonly associated with chronic poverty, according to McKay and Lawson (2002), are being in an unfavorable situation of individuals in their human capital,

demographic composition, location, physical assets, and occupational category, among others. Increasing human capital through education as supported by several studies reduces the probability of individuals to be in chronic poverty. The size of the household and lack of physical assets places additional load on a household's resources. Geographic location also is a key component in the probability of being chronically poor due to the lack of available opportunities to households such as supply of health and education (McKay & Lawson, 2002).

Although some determinants are important to both types of poverty such as human capital, physical assets and demographic composition, some characteristics of transient poverty will still differ from those of chronic since transient poverty is of temporary nature. Some of these are government transfers, seasonality of economic activities and adverse price movements. What empirical evidence suggests is that transient poverty is related with households' unsuccessfully insuring themselves sufficiently against fluctuations in their living conditions leading them temporarily into poverty such as additional member of the household or death of an income earning member (McKay & Lawson, 2002).

The study of Mills and Mykerezi (n.d.) stated that low human capital, minority status and geographic locations are the determinants of poverty in the United States of America (USA) especially in the rural south. Though in most cases, poverty is transient in nature, it should be noted that chronic poverty is a phenomenon especially in low-income economies. Moreover, this is a distinct and irreversible phenomenon that is usually associated with low asset holdings, low income producing activities, disadvantageous demographic characteristics (Mickay & Lawson, 2003). In Latin America, there are cohorts of the population that is likely to be poor compared to others. Those of whom are African descent (which concentrates on minority affiliation), indigenous population, little schooling and households with large member population are considered as the categorical indications of someone in poverty (World Bank, 2003). Like that of in Brazil, educational attainment is likely the cause of poverty but in is more pronounced in household experiencing transient poverty which is explained to be vulnerable to labor market shocks as compared to the resiliency of individuals who attained higher degrees of education (Ribas & Machado, 2007).

Dynamics of poverty and the difference between chronic and transient poor has been discussed by Reyes, et. al. (2011). In their study, the discussion about poverty and the movement of households in and out of poverty was thorough and highly intuitive by examining per capita income and its movement along the poverty threshold. Results showed that there has been a significant decrease in income since from 2003, majority of the transient poor actually has income within the 20 percent of the poverty threshold while in 2006 only one-third of the transient poor fall under this category. The geographical concentration in the Philippines of chronic and transient poor was also examined, wherein results showed that chronic and transient poor are high in Mindanao in which as suggested by Balisican and Fuwa (2004) that welfare of the poor tends to be lower in regions/provinces with political dynasties as compared to provinces with political competition. Moreover, poverty was also examined through the profile of household heads. Male headed households, high school graduates, involved in agricultural sector are only few of the categories mentioned that signifies that a household is more likely in chronic poverty.

The research also observed the profile of household income since income is one of the most notable and quantitative measure of households' capacity to meet daily needs. Income from entrepreneurial activities, especially from agriculture, is the major component of total income of chronic poor while income from non-agricultural sources significantly comprises the income of non-poor. The study further delved on the poor concentrated on agricultural sector wherein the decomposition of households engaged in various agricultural activities and its relation to poverty status are further examined. The study was able to provide a comprehensive and quantitative discussion of poverty in the Philippines.

Panganiban (2010) also delved on the decomposition of poverty using the Cebu Longitudinal Health and Nutrition Survey. The research undertook the similar descriptive process in examining poverty by relating social-economic status to household head characteristics. Significant correlation has been established between poverty, settlement factors, household dependency burden, mother's age, and work in the farming sector. Again, the research is able to point out that poverty heavily exists in the agricultural sector with an emphasis on wage earners as the victims of chronic poverty. The research also considered the time dimension of poverty wherein the possible turnover of socio-economic status from parent to children is subjected to intergenerational analysis. It is theorized that if parents can readily borrow and support investments involving their children (e.g. education), then the mobility in earnings from parents to children would equal the inheritability of endowments, which is not the case in poor families due to restrictions in capacity of financing their investments involving their children. The results of the study showed that there is persistent income status between the two generations and poverty experience can reduce the future income of the child.

These studies suggest that there is concentration and persistent existence of poverty in the agricultural sector. Though this is not an isolated case hence, it is a prominent issue in rural Asia. As pointed out by Balisacan and Fuwa (2003) that the concentration of poverty, both transient and chronic, in Asia is in rural areas and families dependent on agriculture and farming appears to be the poorest. This arrives to policy implications that advancement and changes in the sector will achieve broad based growth.

There are literatures discussed the effect of consumption-smoothing as a way to eliminate or reduce poverty. Kochar (1995) suggested that when a household faces shocks, reallocation of resources to cope with income fluctuations. When crop earnings fails due to natural disasters then farmers can increase their labor income to compensate the lost income. This might also involve reallocation of expenses wherein some households opt to pull their children out of schooling to cope with the loss of income without inferring the long-term implications (Jacoby & Skoufias, 1997).

Poverty Alleviation Programs

Poverty causes inability to afford food, clothing, shelter, education, and health services – all of which are used to measure poverty. The difference between poverty incidences among municipalities and regions will therefore lead to unequal distribution of income and resources among regions and even among the population in a particular region. However, the transmission

mechanisms between the two variables vary. Consider food wherein the relationship from poverty to food inequity is more evident and direct than that of the more complex relationship of food inequity to poverty. A basic and the most explicit relationship between food and poverty are through consumption and income measures. A study conducted by Llanto (1996), delved on the reaction of Philippine households, focusing more on rural and agricultural households, to income and price changes. The motivation for such consideration is due to the belief that households in the lower strata is more affected when there are shocks affecting commodity prices and level of income. Llanto (1996) also highlighted that factors that cause food prices to increase has jeopardizing effects on small farm households hence, sharing the view that "small households spend a relatively large portion of their income on food, and thus, any increase in the price of food products is likely to hurt them more than the richer households". The research also investigated income and price sensitivities with regards to region, income class and location to strengthen the stand of low income and rural households are worse-off when food prices increase and level of household income decreases. It was able to pinpoint that rural households are price inelastic to staples since these are easily accessed by these types of household and are not substitutable.

One transmission mechanism of food disparity to poverty is through agriculture, the primary source of income for the poor. Eradicating food inequality will definitely improve health in the population and there is evidence to suggest that in Africa, gender equality and improved education and health can significantly be a factor in increasing productivity in the agricultural and fisheries sectors of the sub-Saharan Africa (World Bank, 2009). Consequently, increasing these sectors' productivities will lead to the development of these industries creating more employment opportunities for the poor, alleviating them from poverty. Furthermore, with the other case of food disproportion where profits are not fairly distributed among the factors of production, will lead to lower income for the manual laborers who are generally poor pushing more to poverty.

The measure of poverty is not stagnant. It should not be constricted to the lack of income. Its meaning evolves accordingly; from the traditional measure of income as a gauge of individual welfare towards to deprivation of basic capabilities as stated by Sen (1979). Moreover, as the Schelzig (2005) claimed, "poverty is recognized to be a dynamic, complex phenomenon involving concepts such as vulnerability and powerlessness." Furthermore, it is a deprivation of access to other assets that is important for standard living. It is essential to take note of these changes so as to fully represent the state and characteristics of poverty. As discussed by Albert & Molano (2009), in developing countries, poverty lines estimated are absolute poverty lines, which are based on a fixed standard of welfare which is adjusted with regards to price changes. In the Philippines, the estimated poverty line is a representation of income needed to satisfy the minimal needs of a household, both food and non-food. The food aspect is usually referred to as the Food Poverty Line (FPL) which utilizes one-day-menus that meets the required daily dietary needs and nominally valued at the least possible price. On the other hand, the study of Pedro, Candelaria, Velasco and Barba (n.d.) estimated food threshold and poverty incidence using the food baskets across income groups. This is a comparative study between the estimated poverty incidence and food threshold between all income groups versus the bottom 30 percent of the

income group. It showed that food basket of the higher income group consists of food and other commodities that are more complex and expensive as compared to the lower 30 percent of the distribution. Both looked beyond income levels of households. Nutritional intake and food basket composition are taken into consideration to define poverty and individual household welfare.

Since the Aquino Administration in the 1990s, there had been specific projects targeted for poverty reduction. The Social Reform Agenda (SRA) by the Ramos administration focused on poverty alleviation and rural development. The concentration was principally on disadvantaged economic and social groups. This set the foundation for the Social Reform and Poverty Act of 1997 (Republic Act [RA] 8425) that established the National Anti-Poverty Commission (NAPC) that acts as a coordinating and advisory body in programs of social reform and poverty alleviation. It also institutionalizes the "basic sectors" and non-government organizations' participation, support local government units in incorporating SRA, and encourage micro-finance programs and institutions. One recent program launched in 2001 under the supervision of NAPC is the *Kapit-Bisig Laban sa Kahirapan* (KALAHI) program. There are five special projects in the KALAHI program; the rural projects, urban projects, social initiative projects, resettlement areas and in conflict areas.

There are many issues regarding the government's poverty reduction programs basically categorized into policy issues, institutional issues, and resource issues. Under policy issues, every administration is inclined to introduce new programs, usually without concern to what is in progress that was established by preceding administrations. Even successful programs were not continued since they were part of previous presidents' agendas. This results to redundancies in plans, frameworks and targets, and waste of energies and limited resources. Targeting mechanisms were also diverse, inefficient and highly politicized that lead to weak implementation. It also led to inclusion/exclusion of intended beneficiaries and significant leakages to unintended beneficiaries of the programs. Institutional issues, on the other hand, include transitional problems, highly politicized programs and political appointment of agency heads and the representation of the "basic sector" political matters often succeed even from the choice of representatives for the "basic sector", target beneficiaries and the allocation of the budget. The government response for resource issues is that a Poverty Alleviation Fund (PAF) was established in 1998 so that funds for poverty reduction will always be a part of the national budget (Schelzig, 2005)

The extent of research and studies conducted has been substantial and extensive but none of which were able to discuss both the economic and the behavioral reasons why individuals and households stay in or leave out the state of poverty.

CONCEPTUAL FRAMEWORK AND METHODOLOGY

Data Set

In order to quantitatively determine the possibilities and reasons why an individual or a household will stay, depart, and go into the state of poverty, on a national and regional level, the following dataset will be utilized: (1) 2000, 2003, 2006, and 2009 Family Income and

Expenditure Survey (FIES); and (2) 2007 and 2008 Annual Poverty Indicator Survey (APIS). All of which are generated by the NSO. The FIES provides data on family income and expenditure which include among others levels of consumption by item of expenditure as well sources of income in cash and in kind. It specifically discusses levels of living and disparities in income and spending patterns of families belonging to different income groups. Likewise, it also includes related information such as number of family members employed for pay or profit or as wage, salary, or own-account workers; occupation, age, and educational attainment of household head; and other household characteristics. This dataset is released by the NSO every three years. Meanwhile, the APIS is a nationwide sample survey designed to gather information that will provide a comprehensive household socio-economic profile that will provide access and impact indicators useful in the development of an integrated poverty indicator and monitoring system. The APIS is conducted in the years when the FIES is not conducted. Furthermore, it aims to provide accurate, timely, and relevant information for the assessment of poverty alleviation programs and the design of policies intended to reduce poverty.

The sufficient sample of nationwide data contained in the FIES and APIS will allow for the generation of distribution diagrams and measures of living standards in the Philippines for both national and regional level. These measures aim to provide comparable and quantifiable indicators of social welfare that will facilitate interregional comparisons. However, as argued by Jao, Ng & Vicente (2000), since welfare is a multi-facet idea, the attempt to capture its definition into one encompassing indicator remains to be the major limitation of this study. As far as this study is concerned, per capita food consumption expenditure can only serve as a limited proxy indicator of individual welfare. Moreover, because only household level data on consumption expenditure is available, the conversion of household data into per capita consumption expenditure involved some degree of arbitrariness. Although equivalence scales for such conversion are available, they are similarly limited by their inconsistency and subjectivity (Jao, Ng & Vicente, 2000).

Repeated Cross-Section and Pseudo Panels

We will employ the suggested approach of Dang, Lanjouw, Luoto and McKenzie (2011). The suggested procedure will make use of repeated cross-sections that will allow the formulation of pseudo panels to assess bounds of mobility in and out of poverty. Panel data are ideal to measure income mobility but due to data limitation, the authors of the said paper improvised a different approach, which will yield the same results as expected from panel data estimation. It will also offer insightful inferences on state of mobility in poverty and its related dynamics. But it should be noted that the estimates that will be provided are bounds of the fraction of mobility (upper and lower) and not actual point estimates.

The procedure will only make use of FIES 2003 and 2006 because it is conditional that the measure of welfare (consumption or income) will be the same for both periods. FIES 2003 and 2006 utilized the same interview procedure and have the same set of survey questions (Ericta & Fabian, 2009), which make it viable for this estimation. Another issue to consider is that, census surveys are subjected to attrition; there is little probability that a specific household can

be traced from 2003 to 2006. In this regard, the procedure will make use of pseudo panels at cohort level. Hence, both data sets are restricted to households with heads aging from 25 to 60 years old. This restriction will also rationalize that the assessment of mobility in and out of poverty of household with heads aging below 25 and above 60 will be problematic and less indicative since these are the periods that a household might be starting or dissolving.

In estimating the upper bound (or the unobserved first period consumption), the procedure will start with Ordinary Least Squares (OLS) estimation shown by Equation 1.

$$y_{i1}^{1} = \beta_{i1}' x_{i1}^{1} + \varepsilon_{i1} \quad (1)$$

Where:

y_{i1}^{1} is consumption for round 1

x_{i1}^{1} is a vector of household characteristics which are observed in round 1

Predict residuals from Equation 1 and take a random draw with replacement from the distribution. Together with the estimated betas and the observed values of household characteristics in round 2, estimated first round consumption is shown by Equation 2

$$\hat{y}_{i1}^{2} = \hat{\beta}_{i1}' x_{i1}^{2} + \hat{\varepsilon}_{i1}^{2} \quad (2)$$

Where:

$\hat{y}_{i1}^{2}$ is the unobserved first period consumption

x_{i1}^{2} observed household characteristics from round 2

$\hat{\varepsilon}_{i1}^{2}$ randomly drawn betas (with replacement) from (1)

Using the estimated first round consumption, the degree of mobility into and out of poverty will be computed. From Equation 3, hence,

$$Pr\left(\hat{y}_{i1}^{2} < p \middle| y_{i2}^{2} > p\right) \quad (3)$$

Where:

p is the poverty line. The study will use the poverty threshold released by the Philippine National Statistical Coordination Board (NSCB) in 2006, which is PHP 75,285.00[5].

In estimating the lower bounds of mobility, the same procedure will be followed. But instead of the residuals derived from Equation 1 that will be imputed in Equation 2, another OLS will be estimated, shown by Equation 4, and this will replace $\hat{\varepsilon}_{i1}^{2}$ in Equation 2.

$$Pr\left(\hat{y}_{i1}^{2} < p \middle| y_{i2}^{2} > p\right) y_{i2}^{2} = \beta_{i2}' x_{i2}^{2} + \varepsilon_{i2}^{2} \quad (4)$$

Where:

y_{i2}^{2} is consumption for round 2

x_{i2}^2 is a vector of household characteristics which are observed in round 2

The residuals in Equation 4 will serve as the prediction error in Equation 2, which will then provide a way to estimate the lower bound. Once a series of y_{i2}^2 are estimated, movements into and out of poverty of interest will be computed which is the same as the representation in Equation 3.

Discrete Models

Once the bounds of mobility are established, it is also notable to observe the physical characteristics of households and its effect on household mobility into and out of poverty. This approach is used by Abufhele and Puentes (2011) in examining poverty transitions in Chile.

The methodology will use probit and multinomial probit in assessing the factors of transition. These factors center on household characteristics such as age of household head, sex of household head, highest educational attainment of the household head, etc. The probit methodology allows the estimation of probability of entering and exiting poverty. From Equation 5 and Equation 6, the dependent variables in each will take the values of:

$$y_{entered\ t1,t2} \begin{cases} 1 \text{ if household is Non Poor in } t_1 \text{ and Poor in } t_2 \\ 0 \text{ if household is Non Poor in } t_1 \text{ and Non Poor in } t_2 \end{cases} \quad (5)$$

$$y_{exit\ t1,t2} \begin{cases} 1 \text{ if household is Poor in } t_1 \text{ and Non Poor in } t_2 \\ 0 \text{ if household is Poor in } t_1 \text{ and Poor in } t_2 \end{cases} \quad (6)$$

In order to measure the dependent variables, the researchers defined the poor households and the non-poor households with the poverty threshold released by NSCB. Then respective dummy variables are generated in order to specify the conditions of the model. These dependent variables will also be used in the multinomial probit models but there will be two additional dependent variables that will correspond to households that did not change states in between 2003 and 2006. The regression will utilize both household characteristics observed in 2003 and 2006 in order to trace whether factors affecting transition is the same all throughout. The results of this regression will provide clues to policy makers on which aspect of the physical characteristic of households will cause detrimental and/or beneficial effects on whether a particular household will retain or change states.

RESULTS AND DISCUSSIONS

Bounds of Mobility

Due to lack of panel data for Philippine households, this paper utilizes the approach proposed by Lanjouw, Luoto and McKenzie's (2011) in using repeated cross-sections of the household data in estimating the dynamics of poverty by transforming these into a pseudo-panel. Although it did not show how to determine mobility point estimates in the presence of

measurement error, it did present how to measure upper bounds and lower bounds of mobility that may provide some insights to poverty mobility (Lanjouw, Luoto & McKenzie's, 2011).

The estimates in Table 5 are obtained using this approach. Using the upper bound estimates where it assumes no autocorrelation between the 2003 and 2006 error terms, the probability that non-poor households in 2006 were non-poor in 2003 is 85.32 percent and those poor in 2006, the probability that these households were non-poor households in 2003 is 44.34 percent. However, the width of the lower and upper bounds for non-poor households in 2006 is 19.21 percent and 44.34 percent for the poor households in 2006. These very wide gaps of the upper and lower bound estimates limit the insights we can obtain. The most important inference derived from the figures is that movements between states are less often relative to households staying in the same state from 2003 to 2006.

It should be noted here that these numbers are estimated from using age, educational attainment and gender as the only variables to the log of consumption levels. Adding more independent variables would minimize the measurement errors thus narrowing the gap of the bounds (Lanjouw, Luoto & McKenzie's, 2011). Subsequently, the range of mobility will be reduced to where there will be more significant implications.

Table 5
BOUNDS OF MOBILITY

State of the World	Lower Bound	Upper Bound
Non-poor in 2006; Non-poor in 2003	0.6611	0.8532
Non-poor in 2006; Poor in 2003	0.1468	0.3389
Poor in 2006; Non-poor in 2003	-	0.4434
Poor in 2006; Poor in 2003	0.5566	1.0000

Probit Estimations

Table 6 summarizes the marginal effects of various probit estimations. It can be noticed that a particular dependent variable is estimated twice, each considering the household characteristic observed in FIES 2003 and 2006 due to reasons that both FIES are not panels representing the same set of households. Nonetheless, the results will still offer inference regarding the characteristics of the surveyed households and its relation to the transition of being poor to non-poor and vice-versa.

Non-Poor to Non-Poor. Both periods shows the same marginal effects to the probability of staying out of poverty. It is expected that educational attainment is to serve as a key factor in sustaining household security. Furthermore, civil status and in this case, married, is also positive contributing factor to the probability of staying non-poor. With the spouse working it is likely to aid the household state since both household head and spouse are sources of income. Age also contributes positively this maybe due to that most household heads who are considered non-poor in the first period are less likely to retire and/or stop working in the second period.

Poor to Poor. Both periods show the same results when it comes to its response in staying poor. Same with that of Non-poor to Non-poor results, education decreases the chance of

staying poor and the higher educational attainment, the higher the marginal effect of decreasing the probability of staying in poverty. This is also the case in staying married and the spouse of the household head is working. Benefits of being married are likely to decreases chances of being poor since this also offers the opportunity that the spouse will also be working as well providing sustenance to the household.

Table 6
MARGINAL EFFECTS AFTER PROBIT ESTIMATES

	Non-Poor to Poor		Poor to Non-Poor		Poor to Poor		Non-Poor to Non- Poor	
Exogeneous Variables	2003	2006	2003	2006	2003	2006	2003	2006
Age	0.002633	-0.00223	-0.0052	0.00	-0.00238	-0.00441	0.005793	0.007583
Male Household Head	-0.039	0.039814	0.076223	-0.04062	0.024949	0.031403	-0.04726	-0.03131
College Graduate	0.253236	-0.20326	-0.27024	0.341975	-0.13097	-0.14528	0.512482	0.416125
College Undergraduate	0.23922	-0.18097	-0.23743	0.314101	-0.11223	-0.12712	0.460946	0.378754
Highschool Graduate	0.176801	-0.1717	-0.21051	0.259444	-0.10077	-0.122	0.388629	0.338183
Highschool Undergraduate	0.128949	-0.11958	-0.13881	0.203128	-0.06627	-0.07681	0.297921	0.261113
Gradeschool Graduate	0.090875	-0.08498	-0.09706	0.134611	-0.05116	-0.05481	0.224418	0.196181
Gradeschool Undergraduate	0.029589	-0.04308	-0.03749	0.074923	-0.02501	-0.02769	0.099492	0.102318
Married	0.035853	-0.07994	-0.09059	0.050539	-0.02601	-0.04552	0.045842	0.060968
Spouse is Employed	0.033089	-0.03362	-0.06266	0.032239	-0.03309	-0.02902	0.064	0.03287

Non-Poor to Poor. The results for FIES 2003 and 2006 differ in terms of signs. The results for FIES 2003 are counter intuitive, nonetheless, significant. This might be due to the fact that the observed poor households are in 2006 while the non-poor households are in 2003 which causes the response to different period household characteristics to actually contrast. It is plausible that the actual turn over to the new state happens in 2006 and not in 2003 wherein the new state, which is being poor, responded accordingly to a-priori expectations. Observing the figures, it seems results from 2006 makes more sense in this case wherein education of household head diminishes the probability of entering poverty. This is true for all independent variables except for Male Household Head.

Poor to Non-Poor. The same logic applies as that of reasoned in Non-poor to Poor. Education still plays a vital role in fighting poverty. Hence, poor households who readily acquired proper skills and training are more likely to become Non-poor. Furthermore, civil status and if the spouse of household head is working will also likely increase the chances of exiting poverty since it will provide more channels of income to enter a household.

CONCLUSIONS

In the context of this study, we are supposing that government-sponsored programs may have overlooked the root cause of poverty. This is because the persistence of poverty undeniably poses a direct threat toward efficient economic growth and development, more so for the Philippines. This issue has continued to plague the country for decades and as a result, effective poverty alleviation remains one of the main goals for countless administrations.

This time, instead of evaluating government programs aimed to combat poverty, the ideas of transient and chronic poverty would be banked on to demonstrate the behavior of households towards their socio-economic status. In this way, another perspective or approach was considered – the transition of households from being poor to non-poor as to eventually alleviating poverty. This condition in which a household remains within the clasps of poverty can be linked to spending behavior and income changes as well as the effectiveness of government-sponsored programs. Given the nation's state of poverty, we determined the probabilistic relationship of the transition to and from poverty with selected socio-demographic factors. Consequently, we find the data from the FIES as well as the APIS most appropriate as these measures the living standards in the country at the national and regional level. Hence, we employed repeated cross-sections that will allow the formulation of pseudo panels to assess upper and lower bounds of mobility in and out of poverty. Results suggested that the transition to and from poverty from 2003 to 2006 is less probable than remaining in the same state for the same timeframe. Additionally, findings encourage the importance of human capital investment, being married, having an employed spouse, and age in remaining out of poverty as well as escaping from it. Overall, individuals who have acquired proper training and skills are more equipped to sustain their families with a larger income, making them prepared for economic shocks such as financial crises, natural disasters, social conflicts, and environmental property.

As a developing economy besieged by income inequality, inequitable distribution of income, poverty incidence (Todaro & Smith, 2008), and natural calamities, the country's population needs a total overhaul on its human capital development. As evidenced by the results, households headed by educated individuals are more equipped to deal with unexpected shocks that may deter their flow of income and if they do fall into poverty, it is more probable for them to escape from it. As such, securing a stable employment status presents itself as the primary objective. Moreover, the poor must be informed that it is never too late to continue education given technical vocation courses offered by the Technical Education and Skills Development Authority (TESDA) of the Philippines. Instead of providing the most basic skills offered by basic education, TESDA's technical vocation courses provides workers with a skillset more suited for specific jobs. Under-education would also negatively affect spending patterns in which they do not efficiently allocate their resources for long-run benefits. That is, depriving oneself from education alters an individual's behavioral and social factors. The poor have developed a mentality that collectively points to laziness and over-dependency – they do not see the need to take action to solve their own problems because they believe that the state and non-government organizations will eventually provide aid incessantly.

The brisk population growth also emasculates poverty alleviation efforts as the condition of penurious living is passed on across generations. A primary cause of this fast population growth is the inadequate knowledge of individuals on the implications of having a larger household. These households share the common notion that more children they have, the more income can be brought into the household and the higher probability of escaping poverty in the future. These households discount the fact that spending has to be incurred first before these expectations on larger households can be realized. Hence, their assumptions about having a larger household is only true if they can provide basic needs such as education and health for each member of the family.. Failure to provide these basic needs only results to a higher number of dependents in the household, worsening their situation. As population growth is not regulated, the rapid increase in the labor force without basic education will result in high unemployment, which contributes to the persistence of poverty and might be passed on through generations.

The aforementioned concerns are only among countless others that obstruct poverty alleviation in the Philippines. Thus, the economy turns to its administration and government programs in alleviating poverty and achieving equal income distribution. In terms of human capital development, we see progress with the implementation of the K+12 program. However, the government must also exert efforts on providing opportunities for every Filipino to have affordable technical vocational courses by renewing the current process and utilizing information and communication technology (ICT). On the other hand, we also see programs by the government in regulating population growth through the reproductive health (RH) law, which promotes limiting family size through family planning. In general, programs sponsored by the government must be targeted to households who would fully benefit from the program of which must be catered for households suffering from both chronic and transitory poverty. Moreover, more programs such as the conditional cash transfer (CCT) programs must be set in motion as to eliminate the over-dependency of poor households to the government. Past, current, and future administrations need to work together in tweaking implemented programs and in developing new innovations that do not contribute to the redundancy of government programs.

We have been examining why the Philippines is susceptible to extreme cases of poverty. In the pursuit of poverty alleviation, the issue at hand is bringing households out of poverty so that the cycle does not endure across generations. As such, sustained human capital development promotes inclusive growth and is the pathway towards significantly eliminating poverty, attaining equal income distribution. Likewise, adapting the context of America (2013), as in many aspects of life in a developing economy, the response towards poverty is short-term reactive rather than long-term pre-planned. Moreover, the most important response is the attitude an individual takes, the approach undertaken, the organization and institutions working, and most importantly, the commitment made to fight the battle against poverty, as if it were a battle.

ENDNOTES

1 This study was funded by the **Angelo King Institute for Economics and Business Studies** (AKIEBS), De La Salle University (DLSU), Manila, Philippines through the *AKI Research Grants on the Research Thrusts of the School of Economics (2012-2013)*. The authors would like to acknowledge the contributions of Ms. Nashrine B. Aliping and Ms. Mazel G. Pizarro and the research assistance provided by Mr. Dustin Timothy D. Ang. **Disclaimer:** The views expressed in this brief are the views of the authors and do not necessarily represent or reflect the views or policies of the Asian Institute of Management (AIM), the AIM Scientific Research Foundation (AIM SRF), or its Board of Trustees. AIM and AIM SRF make no representation concerning the views in this study and do not assume any legal liability, responsibility nor guarantee the source, originality, accuracy, completeness, or reliability of any statement, information, data, finding, interpretation, advice, opinion, or view presented.

2 John Paolo R. Rivera is the Program Manager of the Asian Institute of Management (AIM) – Dr. Andrew L. Tan Center for Tourism. Prior to this, he was an Assistant Operations Manager at Veritas Financial Partners. He recently served as a consultant at the Economic Research Institute for ASEAN and East Asia. He obtained his PhD Economics, MS Economics, and BS Applied Economics from De La Salle University. He is currently taking the Ateneo-Regis Master in Business Administration at the Ateneo Graduate School of Business. His fields include Quantitative Research Methods, Macroeconomics, International Trade, International Economic Relations, and Development Economics.

3 The income gap is the average income shortfall of the population from the poverty threshold (Todaro & Smith, 2008).

4 The poverty gap is the total income shortfall of the population from the poverty threshold (Todarot & Smith, 2008).

5 As mentioned in the press release of NSCB last March 2008 retrieved from http://www.nscb.gov.ph/press releases/2008/PR-200803-SS2-02_pov.asp

REFERENCES

Abad, R., & Eviota, E. (1983). *Philippine poverty: An annotated bibliography* (Monograph). Makati: Philippine Institute for Development Studies.

Abufhele, A., & Puentes, E. (2011). *Poverty transitions: Evidence for income and multidimensional indicators.* Chile: Department of Economics, Universidad de Chile.

Albert, J., & Ramos, A. (2010). *Trends in household vulnerability* (Discussion Paper Series No. 2010-01). Makati: Philippine Institute for Development Studies.

Albert, J. & Molano, W. (2009). *Estimation of the food poverty line* (Discussion Paper Series No. 2009-14). Makati: Philippine Institute for Development Studies.

Aldaba, F. (2009). *Poverty in the Philippines: Causes, constraints, and opportunities*. Manila: Asian Development Bank.

America, J. (2013, October 26). The Philippines: The most dangerous land on the planet. *GMA News Online.* Retrieved from http://www.gmanetwork.com/news/story/332629/opinion/the-philippines-the-most-dangerous-land-on-the-planet.

Andres, T. & Ilada-Andres, P. (1998). *Effective discipline through Filipino values.* Manila: REX Book Store, Inc.

Balisacan, A.M., & Fuwa, N. (2003). *The Philippine economy: Development, policies, and challenges*. New York: Oxford University Press.

Balisacan, A.M., & Fuwa, N. (2004). Going beyond cross-country averages: growth, inequality and poverty reduction in the Philippines. *World Development*, 32, 1891-907.

Barrientos, A., Hulme, D., & Shepherd, A. (2005). Can social protection tackle chronic poverty? *The European Journal of Development Research*, 17(1), 8-23.

Bekdache, B. (1998). Alternative approaches to modelling time variation in the case of the U.S. real interest rate. *Computational Economics*, 11, 41-51.

Bennett, D. (2008). *The sting of poverty*. Retrieved from http://www.boston.com/bostonglobe/ideas/articles /2008/03/30/the_sting_of_poverty/?page=2

Dang, H., Lanjouw, P., Luoto, J., & McKenzie, D. (2011). *Using repeated cross-sections to explore movements in and out of poverty*. Washington D.C.: World Bank. Retrieved from: https://openknowledge.worldbank.org/ handle/10986/3698

Deininger, K., & Squire, L. (1998). New ways of looking at old issues: inequality and growth. *Journal of Development Economics*, 5: 259–287.

Diebold, F.X., & Rudebusch, G.D.. (1996). Measuring business cycles: A modern perspective. *Review of Economics and Statistics*, 78, 67-77.

Ericta, C., & Fabian, E. (2009). *A documentation of the Philippines' family income and expenditure survey*. Makati: Philippine Institute for Development Studies. Retrieved from: http://dirp4.pids.gov.ph/ris/dps/pidsdps 0918.pdf

Gaiha, R., & Deolalikar, A.B. (1993). Persistent, expected and innate poverty: estimates for semi arid rural south India. *Cambridge Journal of Economics*, 17(4), 409-421.

Garcia, R., & Perron, P. (1996). An analysis of the real interest rate under regime shifts. *Review of Economics and Statistics*, 78, 111-125.

Hamilton, J.D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2), 357-384.

Hine, D., Montiel, C., Cooksey, R., & Lewko, J. (2005). Mental models of poverty in developing nations. *Journal of Cross-Cultural Psychology*, 36(3).

Ilda. (2011). *Ten reasons why Filipinos blame Gloria Arroyo for everything*. Retrieved from http://getreal philippines.com/blog/2011/12/ten-reasons-why-filipinos-blame-gloria-arroyo-for-everything/

Jacoby, H., & Skoufias, E. (1997). Risk, financial markets and human capital in a developing country. *Review of Economic Studies*, 64 (1997): 311-35

Jalan, J., & Ravallion, M. (1998). Transient poverty in postreform Rural China. *Journal of Comparative Economics*, 26(2), 338-357.

Jao, C.S., Ng, C.N., & Vicente, S.C. (2000). *Welfare, distribution and poverty: A national update and interregional Analysis* (Undergraduate Thesis). Manila: De La Salle University.

Kim. (2011). *Top ten why is Philippines poor*. Retrieved from http://topten10most.blogspot.com/2011/05/ top-ten-10-reasons-why-philippines-is.html

Kim, C.J., & Nelson, C.R. (1998). *State-space models with regime switching: Classical and Gibbs-sampling*

approaches with applications. Cambridge: Massachusetts Institute of Technology Press.
Kochar, A. (1995). Explaining household vulnerability to idiosyncratic shocks. *American Economic Review*, 85, 159-64.
Llanto G. (1996). *Philippine household's response to price and income changes* (Working Paper Series 96-05). Makati: Philippine Institute for Development Studies.
Mapa, D. (2011). *State space models and the Kalman filter* (Lecture). Quezon City: School of Statistics, University of the Philippines, Diliman.
McKay, A., & Lawson, D. (2002). *Chronic poverty: a review of current quantitative evidence* (Working Paper No. 15). United Kingdom: Chronic Poverty Research Centre (CPRC).
McKay, A., & Lawson, D. (2003). Assessing the extent and nature of chronic poverty in low income countries: issues and evidence. *World Development*, 31, 425-439
Mills, B., & Mykerezi, E. (n.d.). *Transient and chronic poverty in the US and Rural South* (Manuscript).
Orbeta, A. (2002). *Population and poverty: A review of the links, evidence and implications for the Philippines* (Discussion Paper Series No. 2002-21). Makati: Philippine Institute for Development Studies.
Pagan, A.R., & Schwert, G.W. (1990). Alternative models for conditional stock volatility. *Journal of Econometrics*, 45, 267-290.
Panganiban, M. (2010). *Transient, chronic and integenerational poverty: evidence from the Cebu longitudinal health and nutrition survey*. Quezon City: University of the Philippines
Pedro, M.R., Candelaria, L., Velasco, E., & Barba, C. (n.d.). *Estimating food threshold and poverty incidence using food basket across income groups vs. the bottom 30% income group*. Manila: Asian Development Bank.
Pritchett, L., Suryahadi, A., & Sumarto, S. (2000). *Quantifying vulnerability to poverty: A proposed measure applied to Indonesia* (Working Paper Series No. 2437). Washington D.C.: World Bank.
Ravn, M.O., & Sola, M. (1995). Stylized facts and regime changes: Are prices procyclical? *Journal of Monetary Economics*, 36, 497-526.
Reyes, C.M. (2001). *Poverty profile of the Philippines* (Discussion Paper Series No. 2000-40). Makati: Philippine Institute for Development Studies.
Reyes, C. M., Tabuga, A.D., Mina, C.D., Asis, R.D., & Datu, M.G. (2010). *Are we winning the fight against poverty? An assessment of the poverty situation in the Philippines* (Discussion Paper Series No. 2010-26). Makati: Philippine Institute for Development Studies.
Reyes, C.M., Tabuga, A.D., Mina, C.D., & Asis, R.D. (2011). *Dynamics of poverty in the Philippines: distinguishing the chronic from the transient poor* (Discussion Paper Series No. 2011-31). Makati: Philippine Institute for Development Studies.
Ribas, R.P., & Machado, A.F. (2007). *From transient poverty in Brazil: developing a model for pseudo-panel data* (Working Paper No. 36). New York: International Poverty Centre.
Schelzig, K. (2005). *Poverty in the Philippines: Income, assets, and access.* Manila: Asian Development Bank. Retrieved from http://www.adb.org/documents/books/Poverty-in-the-Philippines/Poverty-in-the-Philippines.pdf .
Sen, A. (1979). *Equality of what?* Retrieved from http://culturability.fondazioneunipolis.org/wp-content/blogs.dir/1/files_mf/1270288635equalityofwhat.pdf.
Sims, C.A., & Zha, T. (2006). Were there regime switches in US monetary policy? *The American Economic Review*, 96, 54-81.
Son, H. (2008). *Has inflation hurt the poor? Regional analysis in the Philippines* (ERD Working Paper No. 112). Manila: Asian Development Bank.
Spears, D. (2010). *Economic decision-making in poverty depletes behavioral control* (CEPS Working Paper No. 213). New Jersey: Princeton University
Todaro, M.P., & Smith, S.C. (2008). *Economic development* (8th edition). Singapore: Pearson South Asia Pte. Ltd.
World Bank. (2003). *Inequality in Latin America and the Carribean – Breaking with history?* Washington, D.C.: The World Bank.
World Bank. (2009). *Gender and Agriculture: Sourcebook*. Washington D.C.: The World Bank, Food and Agriculture Organization (FAO), and International Fund for Agricultural Development (IFAD).

EMPIRICAL INVESTIGATION OF THE RELATIONSHIP BETWEEN LONG TERM INTEREST RATE AND GOVERNMENT DEBT AND DEFICIT SPENDING

Morsheda Hassan, Wiley College

Raja Nassar, Louisiana Tech University

ABSTRACT

In this study, the authors apply time series analysis using data on government spending, debt and long term interest rate in order to determine if government spending or debt has any significant effect on interest rate in the US, England, France, Germany and Japan. In the time series model, 10 year bond rate (a commonly used measure of long term interest rate) was the dependent variable and unemployment rate, inflation rate, GDP and one of the variables (deficit, percent deficit, debt, or percent debt) as the independent variables. Unemployment rate, inflation, and GDP were used as control variables for deficit spending and debt.

On the other hand, unemployment rate and inflation rate were the control variables when percent deficit or percent debt of the GDP was used in the model. Results did show clearly that there was no positive effect of deficit spending, percent deficit, debt, or percent debt on long term interest rate in any of the five countries. These findings lend support to the argument that government borrowing has no crowding effect in the sense that it does not lead to reduced availability of funds for lending and its consequence of higher long term interest rate.

Keywords: Time series analysis, long term interest rate, government debt, deficit spending

INTRODUCTION

Of importance for the economy of a country is the role government spending plays in influencing long term interest rate and hence investment and growth. It is generally recognized that the central bank has a direct influence on short term interest rate, which is the main reason that studies in the literature focus on long term interest rate, as measured by 10-year bond rate, and how it may be affected by government debt and deficit spending. It is argued that debt and deficit spending leads to an increase in long term interest rate which negatively influences investment by the public sector and therefore economic growth. This is explained as being due to reduced availability of funds (caused by government excessive borrowing) for the private sector. However, it is not clear from the literature if debt or deficit spending leads to an increase in long term interest rate and thus curbs investment and economic growth. Hence, more empirical studies are needed in this regard.

How to induce economic growth at the time of a depression or recession is of fundamental importance for the economy of a country. Economists are mixed on this issue. There are those that argue that government borrowing and spending would create jobs and induce economic growth which would lead us out of the recession. Conservative economists on the other hand argue that government spending causes high interest rates which in turn curbs investment and economic growth. Those that argue that government spending induces growth point out that in a depressed economy the private sector is saving more than investing and hence when the government borrows, it is not competing with the private sector for money. Therefore, borrowing under such a circumstance does not cause an increase in interest rate and can only stimulate the economy by increasing employment and demand. Because of the importance of this contrasting economic points of views on policy making by a government, it is essential to determine empirically if government spending has an effect on long term interest rate.

In this study, the authors test this hypothesis by using time series analysis to determine if there is a functional relationship between deficit spending and interest rate in five industrialized countries, namely, the United States, the United Kingdom, Germany, France, and Japan.

RELEVANT LITERATURE

In a study on determinants of long-term interest rate yields in the US, Cebula (2008) using an error correction model reported that there was a possible bidirectional relationship between budget deficit and interest rate yield on tax-free municipal bonds. Also, short term interest rate had a positive effect on long-term rates. Krueger (2003), Cebula (1991), MacAvoy (2003), Friedman (1978), and Meyer (1975) among others have suggested that government budget deficit can lead to what is termed “crowding out” of private investment due to an increase in long-term interest rate. On the other hand, Krugman (2012) argued that government borrowing and budget deficit during an economic downturn or recession has no effect on long-term interest rate and crowding out of private investment. This is so since in a depression the private sector is not borrowing to invest.

Hoelscher (1986), based on an empirical study, reported that government deficits caused long-tem interest rate (measured as yield on 10-year treasury bonds) to rise. An earlier study by Hoelscher (1983) showed no association between budget deficit and long term interest rate for the United States. Other studies in the literature (Making (1983), Motley (1983), Canto and Rapp (1982), Dewald (1983)) reported that deficits had no effect on short-term interest rates as measured by one-year or less treasury bills. Dewald (1983) and Mascaro and Meltzer (1983) reported that there was no relationship between deficit spending and long-term interest rates.

Keith (2005) argued that usually less developed countries show a link between high deficit spending and high inflation. This link, however, was not evident in developed countries like the US. To the extent that high inflation may lead to high interest rate, this implies that in less developed nations a large budget deficit may lead to higher interest rate.

Evans (1985) reported that there was no evidence of a positive association between budget deficit and long term interest rate for the United States. Also, Evans (1987) showed no association for other countries: Canada, France, Germany, Japan, and the United Kingdom.

Pollin (2012) presented evidence which showed that the outstanding debt in the United States tied to the 2009 economic stimulus program had no effect on interest rate . Kiani (2009) showed that there was a significant positive relationship between budget deficit and the slope of the yield curve. Siklos (1988), in an empirical study, found that there was no association between deficit and long term interest rate for Canada.

The literature is mixed on the effect of debts and deficits on interest rates. Hence, more empirical studies are needed in order to shed more light on the situation.

METHODS

Economic data on unemployment rate (Uner), inflation rate (Pinf), debt, deficit, growth domestic product (GDP), and 10- year government bond yield (bond) were collected using the economic data source Gecodia.com for the United States, United Kingdom, Germany, France, and Japan. The data was validated using other data sources where possible.

The different variables for all countries were stationary after first differencing, except for deficit in the US and percent deficit in Japan which did not need to be differenced.

Since there was no significant problem with feedback between the dependent and independent variables, the time series transfer function approach was used where bond was the dependent series and the other variables (Unemployment, inflation, GDP, and debt , percent debt (of the GDP), deficit, or percent deficit (of the GDP) were the independent variables or input series. In addition, the Granger causation test was used in order to determine if any of the independent variables had any significant effect on bond rate.

Time Series Model

A transfer function model between an output series y and input series x_i (i = 1, 2, …,k) is expressed in general as

$$y_t = \sum_i^k v(B)_i \, x_{it} + a_t \qquad (1)$$

Here, $v(B) = \sum v_j B^j$, where B is the backshift operator, $Bx = x_{t-1}$.

The function $v(B)_i$ is determined from the cross correlation between x_{it} and y_t (Wei, 2006).

Once $v(B)_i$ is identified, one can express a_t in Eq. (1) as

$$a_t = y_t - \sum_i^k v(B)_i \, x_{it} \qquad (2)$$

and identify the appropriate model for Eq. (2) from which one can determine the final model in Eq. (1).

In this analysis, Eq. (1) was identified to be of the form

$$y_t = \sum_i^k w_{oi} \, x_{it} + a_t \qquad (3)$$

Our interest here was to estimate the coefficients w_{0i} (i=1,2,..,k) and test for their significance or non-significance in order to determine the effect each x_{it} may have on y_t .

RESULTS

The authors present in the Appendix figures of trends over years for the economic variables (bond rate, percent deficit, percent debt, GDP, unemployment, and inflation) for each of the US, UK, France, Germany, and Japan. It is interesting to see the similarity in trends for the five countries with regard to these variables. This can be attributed to the fact that the economy is global, which leads to interdependence among the economies of different countries.

United States (U.S.)

It is seen from Table 1 that percent inflation is significant with the correct sign, which means that as inflation increases so does the interest rate. Debt was also significant, but had a negative effect, which is contrary to the argument that debt could increase long term interest rate. A negative effect would have no adverse effect on investment and economic growth. GDP and unemployment were not significantly related to bond rate.

Table 1
Parameter Estimates and their Significance for the Model in Eq. (3) for the U.S.

Variable	Parameter	Estimate	t-value	Pr
Uner	w_{01}	0.06299	1.26	0.2167
Pinf	w_{02}	0.223	3.33	0.0020
GDP	w_{03}	-.0008446	-1.24	.2246
Debt	w_{04}	- 0.0009235	-2.12	0.0409

When Gdp was dropped from the equation, the debt estimate became -0.0007875 with a probability of 0.0726, not significant at the 0.05 level. The sign of the estimate is again negative. Results with regard to pinf and debt did not change when, in addition to Gdp, Uner were dropped from the model. It is seen from Table 2 that deficit had no significant effect on bond. Only inflation had a significant positive effect as seen from the parameter estimate and the p- value of 0.0054.

Table 2
Parameter Estimates and their Significance for the Model in Eq. (3) for the U. S.

Variable	Parameter	Estimate	t-value	Pr	Lag
Uner	w_{01}	0.03178	0.68	0.5010	0
Pinf	w_{02}	0.20931	2.96	0.0054	0
GDP	w_{03}	-0.0014521	-1.53	0.1340	0
Debt	w_{04}	0.0015637	1.43	0.1607	0

The results were the same when GDP and Uner were deleted from the model for not being significant. In this case deficit was not significant (Pr=.5370) and Pinf remained significant with Pr = .0054 and a w_o estimate of 0.07307. These above results agree with the Granger causation test where deficit had no significant relationship with bond rate (pr = 0.5985). Also debt had no significant effect on bond rate (pr= 0.4851).

Percent Deficit and Percent Debt

When percent deficit was used in the model in place of deficit, the parameter estimate was not significant (0.03071, pr = 0.7191). Also, when percent debt was used instead of debt, results showed that it had negative effect on bond rate, but not significant at the 0.05 level (-.08314, pr =0.0755). The Granger causation test also showed that there was no significant relationship between bond rate and percent deficit or percent debt (pr = 0.5 and 0.26, respectively).

United Kingdom (UK)

Results in Table 3 show that debt had a significant negative effect on bond rate which is contrary to the argument that debt can increase long term interest rate. Both GDP and unemployment had no significant effect on bond rate.

Table 3
Parameter Estimates and their Significance for the Model in Eq. (3) for the UK

Variable	Parameter	Estimate	t-value	Pr
Uner	w_{01}	-0.12811	-0.51	0.6101
Pinf	w_{02}	-0.01590	-0.59	0.5579
GDP	w_{03}	-0.02953	-2.11	0.0415

When Gdp and Uner were deleted from the model, debt was still significant with W_{03} = -0.03254 (pr = 0.0164). The Granger test did not show any significant relationship between Bond rate and debt (pr = 0.3425). Table 4 gives the estimates of the model when deficit is used. Deficit is not significant, showing no relationship to bond rate. Also, the Granger test did not show any relationship between deficit and bond rate (pr = 0.5393).

Table 4
Parameter Estimates and their Significance for the Model in Eq. (3) for the UK

Variable	Parameter	Estimate	t-value	Pr
Uner	w_{01}	-0.30887	-1.07	0.2918
GDP	w_{02}	-0.02895	-1.03	0.3091
Deficit	w_{03}	-0.01580	-0.41	0.6822

Percent Deficit and Percent Debt

Percent deficit, when used instead of deficit, had no significant effect on bond rate (-0.1074, pr = 0.217) . Also, percent debt had no significant effect on bond rate (-.0539, pr = 0.159). These results were in agreement with the Granger causation test which showed no significant relationship between bond rate and percent deficit or percent debt (pr = 0.26 and 0.107, respectively).

Germany

Table 5 shows that only Pinf had a significant positive effect on bond rate (pr = 0.0001). Debt, in particular, had no effect on bond rate. When Uner and Gdp were deleted from the model, Debt still had no significant effect on bond rate (W_{04} = -.00796, pr = .2047) while the effect of pinf was still significant (W_{02} = 0.5415, pr = 0.0001).

Table 5
Parameter Estimates and their Significance for the Model in Eq. (3) for Germany

Variable	Parameter	Estimate	t-value	Pr
Uner	w_{01}	-0.0057240	-0.04	0.9651
Pinf	w_{02}	0.52118	4.27	<0.0001
GDP	w_{03}	0.0055773	0.48	0.6372
Debt	w_{04}	-0.01053	-1.20	0.2386

It is seen from Table 6 that only percent inflation had a significant effect on bond rate. The deficit effect was not significant at the 0.05 level. When Uner and Gdp were deleted from the model, results showed that inflation was still significant (W_{02} = 0.5446, pr = <0.0001) and deficit not significant (W_{03} = 0.0107, pr = 0.1368) . The Granger test showed that there was no significant relationship between deficit and bond rate (pr= 0.1320) or debt and bond rate (pr = 0.2051).

Table 6
Parameter Estimates and their Significance for the Model in Eq. (3) for Germany

Variable	Parameter	Estimate	t-value	Pr > |t|
Uner	w_{01}	0.08586	0.62	0.5407
Pinf	w_{02}	0.51710	4.38	<0.0001
GDP	w_{03}	0.01328	1.02	0.3154
Deficit	w_{04}	0.02028	1.73	0.0923

Percent Deficit and Percent Debt

Analysis using percent deficit and percent debt instead of deficit and debt showed that there was no effect of either one on long term interest rate. Estimates of coefficients for percent deficit and percent debt were 0.044 (pr = 0.2479) and -0.0456 (pr = 0.3201). The Granger test showed a

relationship between bond rate and percent deficit as well as percent debt. However, these relationships were not quite significant at the 0.05 level (pr = 0.075 for percent deficit and 0.065 for percent debt).

France

Results for France (Table 7) showed that inflation had a positive and significant effect on bond rate. Debt was almost significant at the 0.05 level. However, its effect on bond rate was negative rather than positive, which is contrary to the argument in the literature.

Table 7
Parameter Estimates and their Significance for the Model in Eq. (3) for France

Variable	Parameter	Estimate	t-value	Pr > |t|
Uner	w_{01}	0.2849	1.10	0.2792
Pinf	w_{02}	0.5331	4.26	0.0002
GDP	w_{03}	0.02158	1.02	0.3135
Debt	w_{04}	-0.02176	-1.91	0.0656

When GDP and Uner were deleted from the model, results showed that inflation was still significant (W_{02} = 0.389, pr < 0.0001), but not debt (W_{04} = -.01399, pr = 0.1263). Table 8 shows that only Pinf was significant and had a positive effect on bond rate. Deficit spending had no effect on bond rate.

When the model was reduced by deleting Uner and Gdp, results still showed that pinf was significant (W_{02} = 0.6326, pr <0.0001), but not deficit (W_{04} = -0.0423, pr = 0.2477). Granger analysis showed that there was no significant relationship between deficit and bond rate (pr = 0.2591) or between debt and bond rate (pr = 0.3532).

Table 8
Parameter Estimates and their Significance for the Model in Eq. (3) for France

Variable	Parameter	Estimate	t-value	Pr > |t|
Uner	w_{01}	-0.14042	-0.51	0.6133
Pinf	w_{02}	0.62751	4.37	0.0002
GDP	w_{03}	0.0032236	0.15	0.8841
Deficit	w_{04}	-0.05629	-1.29	0.2090

Percent Deficit and Percent Debt

Percent deficit, when included in the model, had no significant effect on bond rate (-0.230, pr = 0.159). However, percent debt was almost significant at the .05 level (-0.08843, pr = 0.0544), but had a negative sign, which is contrary to the argument that debt could increase long term interest rate. The Granger test results showed that both percent deficit and percent debt were not related to bond rate (pr = 0.164 and 0.175, respectively).

Japan

It is seen from the probability column in Table 9 that none of the independent variables is significantly related to bond rate. When Uner, Pinf, and Gdp were dropped from the model, results showed a negative effect of debt on bond rate which was significant with probability 0.0667 (W_{04} = -0.007223, pr = 0.0667). This was, however, the wrong sign in the sense that increased deficit causes a reduction rather than an increase in long term interest rate. On the other hand, the Granger causation test showed no relationship between bond rate and debt (pr = 0.597).

Table 9
Parameter Estimates and their Significance for the Model in Eq. (3) for Japan

Variable	Parameter	Estimate	t-value	Pr > |t|
Uner	w_{01}	-0.17932	-0.44	0.1587
Pinf	w_{02}	0.05436	1.44	0.6413
GDP	w_{03}	0.0036737	0.47	0.6413
Debt	w_{04}	-0.0058186	-1.28	0.2081

Results in Table 10 show that there was no significant effect of deficit on bond rate. Also, unemployment, inflation, and GDP had no effect on bond rate. When unemployment, inflation, and Gdp were dropped from the model, deficit showed an effect on bond rate, but was not significant at the 0.05 level (W_{04} = 0.01893, pr = 0.088). This effect, however, is negative since deficit as used in the analysis was negative in value. A negative effect implies that as deficit increases bond rate decreases, which is contrary to the argument in the literature. The Granger causation test did not show any significant effect of deficit on bond rate (pr = 0.7991).

Table 10
Parameter Estimates and their Significance for the Model in Eq. (3) for Japan

Variable	Parameter	Estimate	t-value	Pr > |t|
Uner	w_{01}	-0.25536	-0.67	0.5058
Pinf	w_{02}	0.05507	1.46	0.1538
GDP	w_{03}	-0.0002957	-0.04	0.9682
Deficit	w_{04}	0.01348	1.14	0.2630

Percent Deficit and Percent Debt

Analysis using percent deficit and percent debt in the time series model showed that there was no significant effect of either variable on bond rate. Estimates of coefficients were -.0228 (pr = 0.275) for percent debt and 0.0261 (pr = 0.315) for percent deficit. Also, the Granger test did not show any significant relationship between bond rate and percent deficit (pr = 0.553) or percent debt (pr = 0.655).

DISCUSSION AND CONCLUSION

Results of the time series analysis and the Granger causation test for the 5 countries showed an overwhelming support for the argument that debt and deficit have no effect on increasing the long term interest rate. It is seen that when debt or deficit had a significant effect on long term interest rate it was to decrease the rate rather than to increase it. In all the results, except for Japan, there was as expected, a positive and significant effect of inflation on bond rate.

The results of this study do not lend support to the argument that increased debt or deficit through borrowing by the government leads to an increase in interest rate which can have an adverse effect on economic growth by curbing investment by the private sector. These results seem to imply that when governments borrow, especially in times of economic recessions or slow down, they do not seem to compete for money with the private sector and hence do not cause any shortage of money which can lead to an increase in interest rate which can curb investment and economic growth. This is in essence the arguments of many Kensyian economists, especially during recession time. Our results support this argument and could mean that government debt and deficit spending, at least at the level seen in the present data, can stimulate economic growth and do not lead to an increase in interest rate and reduced investment by the private sector.

Inflation in this analysis is, as expected, positively related to bond rate. The few cases in this analysis which indicated that debt or government spending has a negative effect on long term interest rate, could be due to an indirect effect related to the central bank outlook on inflation.

The present analysis is with regard to five industrialized countries with relatively large and strong economies. This could be a reason why deficit spending and debt have no influence on long term interest rate. In smaller countries with smaller economies excess government borrowing, especially in a currency not of their own, can be expensive in that it can lead to higher interest rates due to the fact that investors demand higher rates because they may fear government default on their investment. In the future, it would be of interest to study the relationship between government spending and long term interest rate for such countries.

REFERENCES

Canto, V. A., & Rapp, D. (1982). The crowding out controversy. Arguments and Evidence, *Federal Reserve Bank of Atlanta Economic Review*, 67, 33-37.

Cebula, R. J. (2008). Determinants of long-term interest rate yields: The case of the US. ICFAI Journal of Applied Economics, 7 (3), 37-49, The ICfai University Press.

Cebula, R. J. (1991). A Note on Federal Budget Deficits and the Term Structure of Real Interest Rates in the United States, *Southern Economic Journal*, 57, 1170-1173.

Dewald, W. G. (1983). Federal deficits and real interest rates: Theory and evidence Federal *Reserve Bank of Atlanta Economic Review*, 68, 20-29.

Evans, P. (1985). Do large deficits produce high interest rates? *American Economic Review*, *7*, 68–87.

Evans, P. (1987). Do budget deficits raise nominal interest rates? *Journal of Monetary Economics*, *20*, 281–300.

Friedman, B.M. (1978). Cowding out or crowding in? Economic consequences of financing government deficits. Brooking papers on economic activity, 3, 593-641.

Hoelscher, G. (1983). New evidence on deficits and interest rates. *Southern Economic Journal*, *50*, 319–333.

Hoelscher, G. (1986). New evidence on deficits and interest rates. Journal of Money, Credit and Banking, 18 (1), 1-17. The Ohio STtae University Press.

Keith, S. (2005). Do Budget Deficits Cause Inflation? Philadelphia Fed., www.philadelphiafed.org

Kiani, K. M. (2009). Federal Budget Deficits and Long-Term Interest Rates in USA *The Quarterly Review of Economics and Finance*, 49, 74–84

Krueger Alan B (2003), "Big Federal Deficits, Bigger Risks", New York Times, March 6.

Krugman, P. 2012. End This Depression Now. W.W. Norton and Company, New York, London.

MacAvoy P. (2003), "'Don't Just Stand There...' Treasury Secretary William E Simon and Fiscal Policy During Stagflation 1975-76", *Atlantic Economic Journal*, 31, 213-218.

Making, J. H. (1983). Real interest, money surprises anticipated inflation, and private sector behavior. *American Economic Review*, 73, 994-1010.

Mascaro, A. and A. H. Meltzer. (1983). Long-and short-term interest rates in a risky world. *Journal of Monetary Economics*, 12, 485-518

Meyer, L.H. (1975). The balance sheet identity, the government Financing constraints, and the crowding-out effect. *Journal of Monetary Economics*, 1, 65-78.

Motely, B. (1983). Real interest rates, Money and government deficits. *Federal Reserve Bank of San Francisco Economic Review*, 31-45.

Pollin, R. (2012). US government deficits and debt amid the great recession: what the evidence shows. *Cambridge Journal of Economics* , 36, 161–187.

Siklos, P. L. (1988). The deficit-interest rate link: empirical evidence for Canada. *Applied Economics*, 20, 1563-1577.

Wei, W. (2006). Time Series Analysis: Univariate and Multivariate Methods. Pearson Addison Wesley, New York

APPENDIX

Figure 1
Trends over years of bond rate for the US (----), Japan (----), Germany (----), France (----), and UK (----).

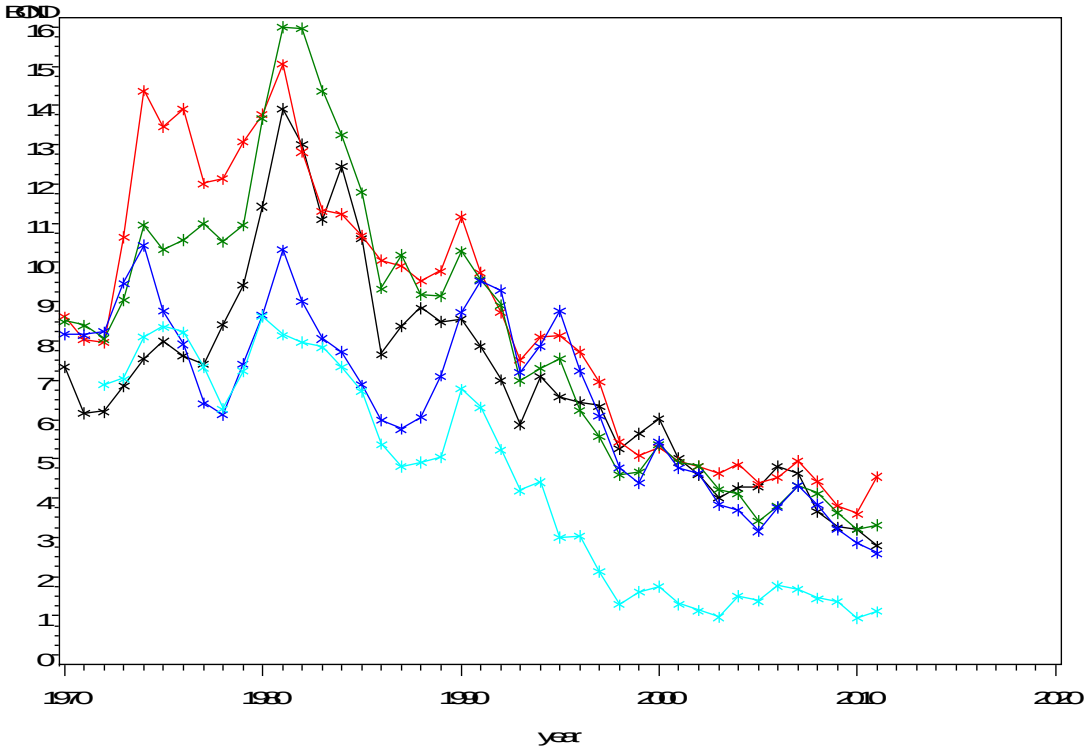


Figure 2
Trends over Years of Percent Deficit for the US (----), Japan (----), Germany (----), France (----), and UK (----).

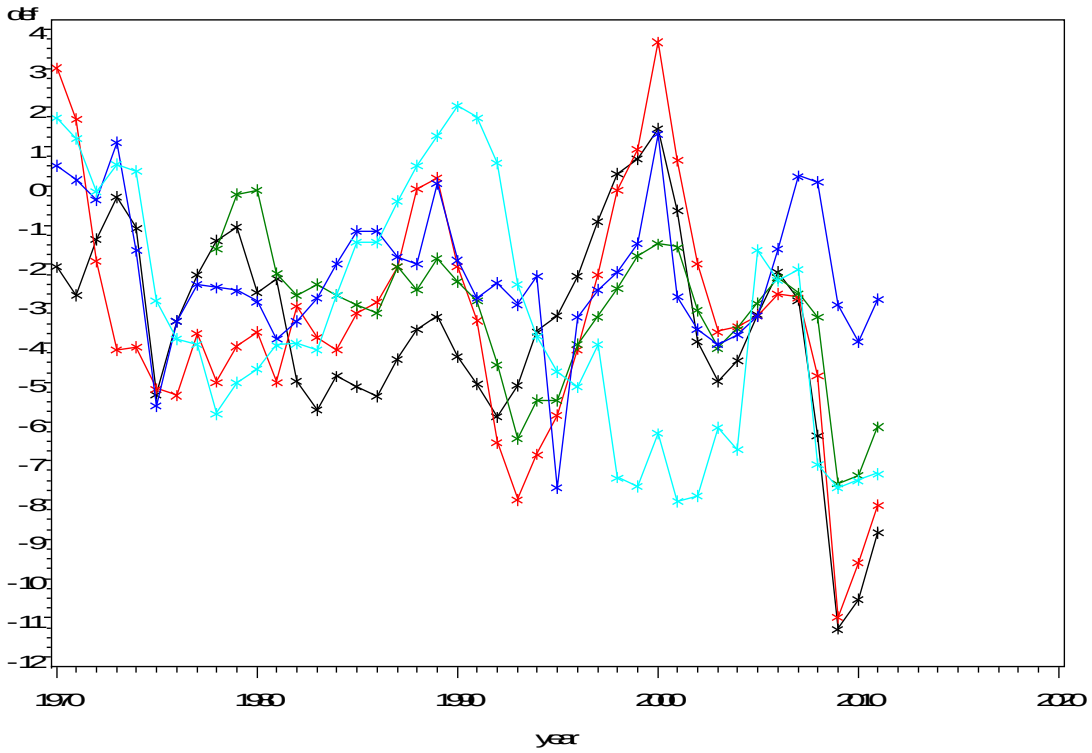


Figure 3
Trends over Years of Percent Debt for the US (----), Japan (----), Germany (----), France (----), and UK (----)

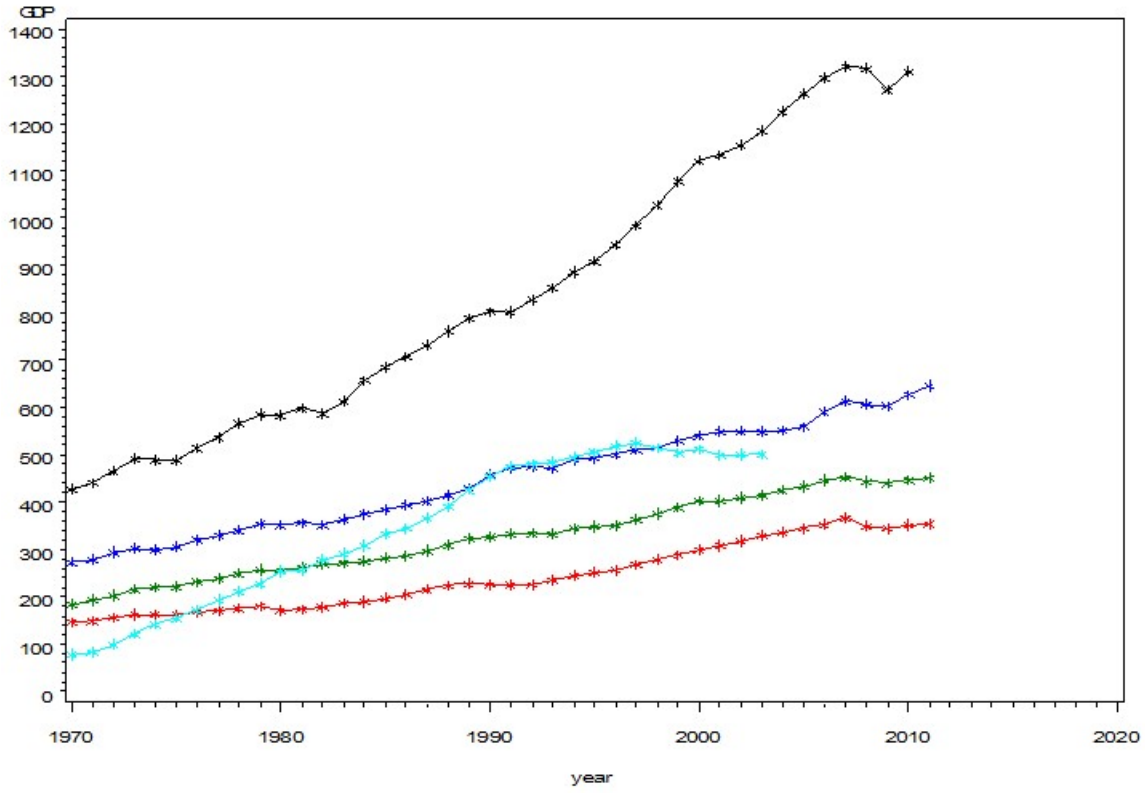


Figure 4
Trends over Years of GDP for the US (----), Japan (----), Germany (----), France (----), and UK (----). GDP for the US is in 10 Billion Dollars, Japan in jpy tr, UK in Billion Bunds, France in Billion Euros, and Germany in Billion Euros

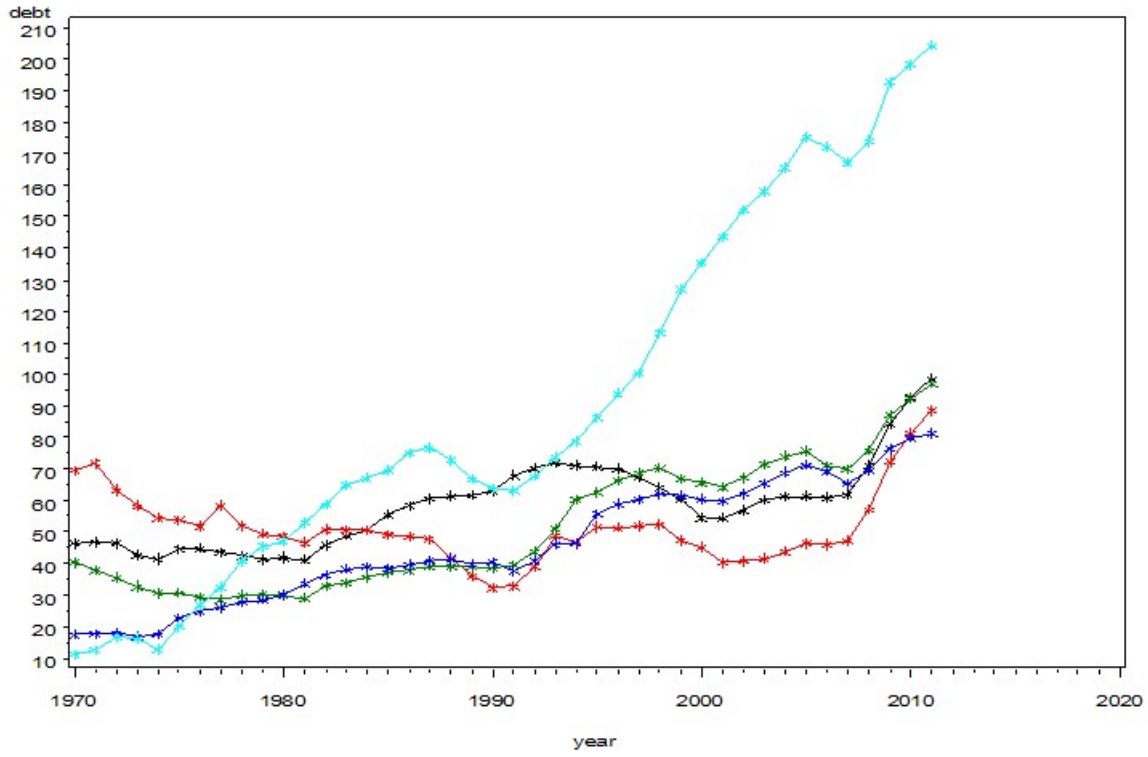


Figure 5
Trends over Years of Inflation Rate for the US (----), Japan (----), Germany (----), France (----), and UK (----)

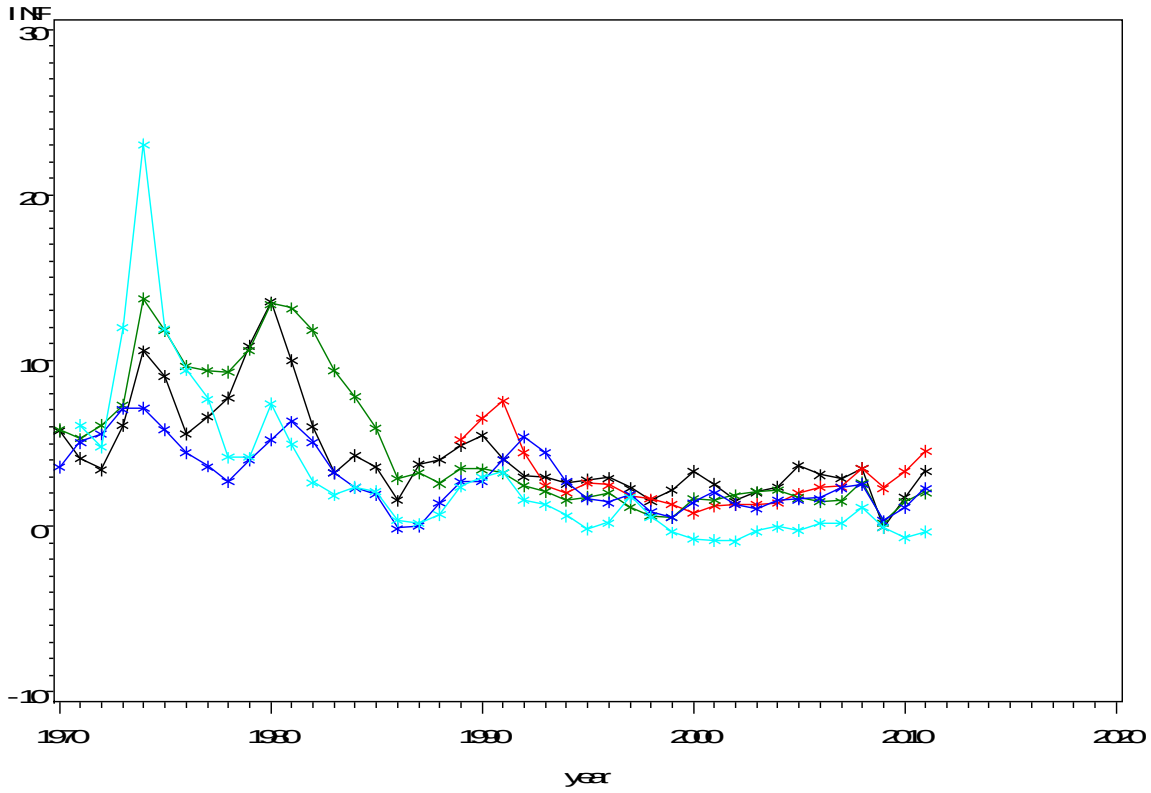
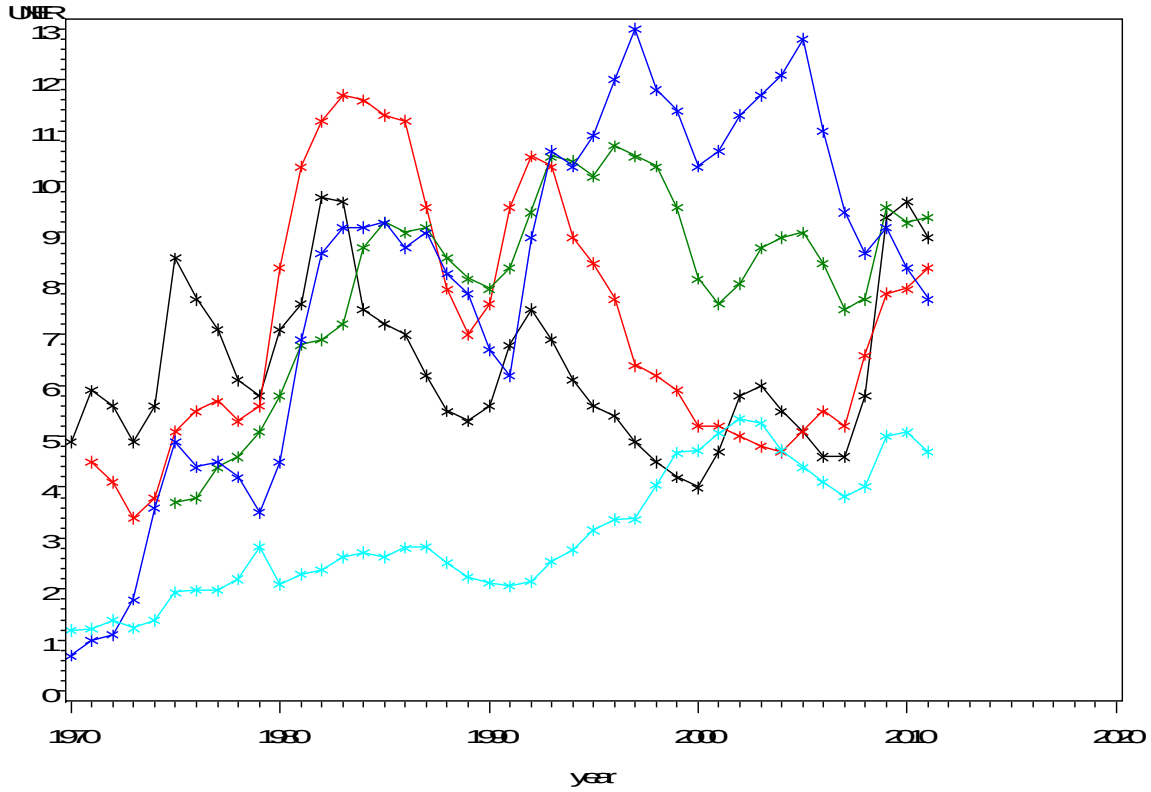


Figure 6
Trends over Years of Unemployment Rate for the US (----), Japan (----), Germany (----), France (----), and UK (----)



REGIONAL DIFFERENTIAL WEALTH EFFECT ON HOME VALUE: A CROSS-SECTIONAL ANALYSIS

Askar H. Choudhury, Illinois State University
Nathan S. Hartman, Illinois State University

ABSTRACT

Home prices are regarded as an important determinant for consumption and considered to be a leading indicator for the economy. However, homebuyers' evolving preferences on new or existing characteristics can have different impacts on housing price for different homebuyers differently. Thus, regional home price differences may be due to differential effect in homebuyers' preference that is reliant on individuals' wealth barometer. In general, housing market in a region with higher median home values would also have a higher household income. This gives credence to the proposition that region with higher percentage of high income earners are instrumental for elevated home price. The unequal distribution of economic pulsation is in part dependent on the concentration of wealth that is different in different geographic locations. It is expected that housing prices are determined by income and therefore wealth, but it is not certain that there is a differential effect on home value due to different regions. Thus, the objective of this study is to understand the dynamic relationships of the wealth effect and regional effect on the home value. In particular, statistically significant relationship between home value and wealth factor dissimilarity due to locational differences are observed. Moreover, after controlling for population size and unemployment rate, wealth factor impacts the value of the house positively across the regions.

INTRODUCTION

Housing market literature has presented a variety of housing characteristics to explain home value. Regression analysis is typically applied to identify the direction of the relationship between home value and these housing characteristics. Predicting home value can be daunting as homes in general contain countless physical attributes that influence their values and can include lot size, square footage, number of bedrooms, number of bathrooms, and other characteristics (Zietz, Zietz, & Sirmans, 2008). Popularity for this avenue of study has been highlighted by Sirman et al.'s (2005) review of studies that listed age of home and square footage to be the structural features frequently found to affect the selling price of a home. Amenities such as fireplaces, central air-conditioning, and patios were also found to influence home price positively. Neighborhoods' attributes both natural and manmade were also major components for home valuations. School districts, access to local parks and recreations have positive impacts on housing prices (Clark & Herrin, 2000; Downes & Zabel, 2002; Figlio & Lucas, 2004; Reback, 2005; Seo & Simons, 2009; Garcia, Montolio, & Raya, 2010). However, homebuyers' evolving

preferences for particular characteristics continues and thus new or existing characteristics can have different impacts on housing price at different points in time and for different homebuyers differently (Malpezzi, 2003). As a result, home characteristic studies continue to be published to maintain a contemporary understanding of predictors' behavior on the home price. Other additions to the home value literature consider how differences in regional home price due to different geographic locations can create a differential effect in homebuyers' preference.

Home prices are regarded as an important determinant for consumption (Case, Quigley, & Shiller, 2005) and considered a leading indicator for the health of the economy. This is a serviceable surrogate variable for understanding the immensity of consumerism in the U.S. because trends in home price are a simple tool to infer economic health by city, state, region, and country. Other types of buying decisions made by individual consumers are less desirable indicators because the data is more difficult to capture as choices are almost limitless in number and occur for anyone homeowner in a variety of locations. Moreover, frequent change in consumer buying trend makes buying behavior for most consumer goods an unreliable representation of economic health across different geographic regions and time periods. In general, macroeconomic variables determining housing price on a regional scale are preferable for those seeking determinants of the U.S. economy's health, while microeconomic variables are preferable to those trying to explain local home prices and local real estate markets. While macroeconomic studies are less common, they can be a crucial linkage to understand the health of the economy.

Researchers have proposed many different models to predict the value of housing across multiple studies, where the relationship between housing price and other variables were separately analyzed by geographic locations, including national, regional, state, and metropolitan areas (Case et al., 2005; Zhu, Füss, & Rottke, 2013; Fullerton, Fierro, & Donjuan-Callejo, 2009; Fullerton & Villalobos, 2011; Holmes, Otero, & Panagiotidis, 2011). Sirmans et al. (2005) suggested the positive effect square footage had on selling price was similar in some of the multi-state geographic regions (Census Bureau designated regions of the U.S.: Northeast, Midwest, Southeast, and West). They also discussed that lot size, fireplace, central air-conditioning, basement, garage, and perceived school quality have a positive relationship with housing price across all regions.

Home price in this literature has represented as an approximation of a home's actual value, and on a larger scale the median home price by city, state, or region is representative of the overall home value within that market. Changes in income and consumer sentiment are the major determinants for the fluctuations in housing price (Boelhouwer, Haffner, Neuteboom, & de Vries, 2004; Rouwendal & Longhi, 2008). Consumption has been found to respond to changes in home prices, but consumption effects were weaker for households with unused borrowing capacity (Campbell & Cocco, 2007). A subsequent study by Algieri (2013) found housing prices were most sensitive over the long term to income changes, population, and inflation rate. Riley (2012) observed that among all types of homes, expensive houses exhibited more consistent price appreciation and strong relations to changes in regional home values. This

gives credibility to the proposition that regions with higher percentage of high income earners experience greater home price appreciation and have a stronger regional economy. These relationships were confirmed, in part, because homes were clustered together in locations, states, and regions by home value. As an example, homes with lower median price are concentrated in areas that are geographically separated from homes with higher median prices. Therefore, we expect regions at a lower median home price to contribute less to the GDP than regions with higher median home price. This variable (median home price) distills regional spending habits and wealth accumulation. Similarly, consumer spending habits and wealth are strongly correlated with home values. In other words, markets (local and regional) with higher median home values would also have a higher household income. This effect has been observed on a localized scale where spending habits are different in low income neighborhoods (Sirmans & Macpherson, 2003; Feldman, 2002), with lower valued homes that tend to have more fast food (low priced) restaurants (Block, Scribner, & DeSalvo, 2004). In addition, lower income neighborhoods have more liquor stores as observed by LaVeist and Wallace (2000), while Morland et al. (2002) found wealthier neighborhoods have more supermarkets and more "heart-healthy" foods. To this end, Kenny and Reinke (2011) found that educational attainment, household income, and home values predicted neighborhood wealth.

Government intervention encourages individuals to value home ownership as a way to improve local, regional, and national economy, which suggests a need for better understanding of home price predictors. Specifically, government incentives (such as, mortgage debt tax deductions) were designed to encourage homeownership (Poterbal & Sinai, 2011) as a proactive policy built on the premise that growth in the rate of homeownership also increases the wealth of the nation. In addition, the negative impacts of short-term regional economic instability would have less impact on the national economy as the homeownership incentive encourages homeowners to retain homes and promote stability in communities and financial markets (Rosenthal, 2008). Homeownership accounts for a sizeable number of jobs regionally in terms of maintenance, modification, and insurance, but also nationally through the homebuilding, realty, and financial industries. In other words, these incentives attempt to weaken the relationship between disruptions in short-term homeowner income and home values. Campbell and Sances (2013) observed that at the state level decline in home value and personal income impacts state's budget greatly. During the recent recession, budget gaps were larger in states where personal income and home values declined the most, and this resulted in altered spending behavior, which in turn drove up the demand for social services and thus created larger fiscal gaps. Thus, the relationship between spending behavior and home values connection to tax revenues highlight another interconnected path of home prices' to the economy. In general, per capita income, unemployment, and interest rates are key determinants for short-term fluctuations in consumer spending on housing. Thus, fluctuations in these factors also create changes in housing prices. Therefore, when income increase, so do property prices and its value; conversely when earnings drop, so do housing prices and its value. For most US households, the home is the most important asset (Davidoff, 2003). However, the results of policies encouraging homeownership

incentives are far from ideal as the benefit to low-income households has been inconsistent (Duda & Belsky, 2002; Turner & Smith, 2009; Riley, 2012). On a national scale, locations where low-income homeowners are heavily concentrated experience perpetually low median home values and weaker home price appreciation compared to high-income locations.

The unequal distribution of economic pulsation is in part dependent on the concentration of human capital in different geographic locations. These concentrations of human capital also tend to reflect similar wealth concentrations in those regions. Kenny and Reinke (2011) found populations with more human capital, in terms of the population's level of education attainment, spent more on housing and likely put upward pressure on home price in those locations. While human capital is an important predictor of eventual earning power and purchasing power, such that a larger concentration of educated individuals would increase buying power and thus increase home prices, may also construct a confounding effect of education with wealth on the home value. Nonetheless, high income earners disproportionately drive housing prices upward, because of their access to better credit and less income/employment volatility. Therefore, relatively richer neighborhoods are more likely to become a new part of a city than poorer neighborhoods (Kenny & Reinke, 2011) and this increases home value as others aspire to live in these locations and creates a lasting effect. Supporting research has found higher income earners have more predictable homeownership behavior, as 85% of high-income households have homeownership. While, low-income households only experience a 50% homeownership rate (Turner & Smith, 2009). This further demonstrates characteristics of specific sub-populations that are effective in determining the regional fluctuations in home values. The primary implication of these variations is that the impact of the wealth effect is not equal for all regions, and therefore, policymakers should adjust and control relevant factors to understand the differential wealth effect.

Thus, the objective of this study is to understand the dynamic relationships of the wealth effect and regional effect on the home value. In particular, the relationship between home value and wealth factor dissimilarity due to locational differences warrant our attention. Based on the typical research study, it may not be surprising to anticipate that housing prices are determined by income and therefore wealth, but it is not certain that home value can have a differential effect due to regional differences. Therefore, undertaking this research analysis may help determine the dynamic relationship between these factors. For that reason, we first examine the relationship between home value and wealth factor for each region. This will provide us the general direction of the relationship between home value and wealth. Knowing the direction of this relationship, we then combine the regions together in one model to understand the comparative differential effect of wealth due to the different regions considered in this paper. This will enable investors and policymakers to determine how home value responds to the wealth disparity depending on the region they belong.

DATA DESCRIPTION AND RESEARCH METHODOLOGY

For our analysis, we have obtained data from the US Census. The data includes "median home value" for each fifty states among other factors (discussed below) for the year 2000. In order to determine the effect on home value, we have identified and applied several socio-economic factors as independent variables in the models. We have also considered regional effect on the home value as the regions are defined by the US Census (see, regions descriptions below). All the factors used for this research analysis were obtained from the same source to maintain the data reliability and consistency. This is the only year (year 2000) that US Census provided all these various socio-economic factors in their report.

Regions

As mentioned earlier, data will be divided and analyzed into four different regions.

List of States by Region			
Northeast Region (Region 1)	**Midwest Region** (Region 2)	**South Region** (Region 3)	**West Region** (Region 4)
Maine	Ohio	Delaware	Montana
New Hampshire	Illinois	Maryland	Idaho
Vermont	Indiana	Virginia	Wyoming
Massachusetts	Michigan	West Virginia	Colorado
Rode Island	Wisconsin	North Carolina	New Mexico
Connecticut	Minnesota	South Carolina	Arizona
New York	Iowa	Georgia	Utah
New Jersey	Missouri	Florida	Nevada
Pennsylvania	North Dakota	Kentucky	Washington
	South Dakota	Tennessee	Oregon
	Nebraska	Alabama	California
	Kansa	Mississippi	Alaska
		Arkansas	Hawaii
		Louisiana	
		Oklahoma	
		Texas	

To explore the relationship with median home value (in US dollars), we concentrated only on the following socio-economic factors: percentage of individuals with bachelor's degree or higher, the unemployment rate, the total population, median earnings of male, median earnings of female, and the percentage of households with income greater than $100,000 (which is considered as wealth indicator). In addition, we have also explored other factors, such as, GDP, per capita income, rental cost, monthly maintenance expenses, etc. in our preliminary analysis (results not shown). The model will be constructed using these above mentioned socio-

economic factors with the data obtained from US Census of year 2000. The analysis will be done region-wise as the regions are determined by US Census based on their geographic locations and also a combined analysis for all regions together.

TABLE-1A: Summary Statistics of Region 1 (Northeast).					
Variables	**N**	**Mean**	**Std Dev**	**Minimum**	**Maximum**
Home value (Median)	9	138400	32160	97000	185700
BA + %	9	27.86667	3.66776	22.40000	33.20000
Unemployment Rate	9	5.21111	0.99177	3.80000	7.10000
Earnings-M	9	39399	5122	32372	46368
Earnings-F	9	28963	3454	24251	33318
Wealth (Income_G_100K)	9	13.98889	5.03474	7.10000	21.30000
Population	9	5954931	6315772	608827	18976457

TABLE-1B: Summary Statistics of Region 2 (Midwest).					
Variables	**N**	**Mean**	**Std Dev**	**Minimum**	**Maximum**
Home value (Median)	12	98075	18440	74400	130800
BA + %	12	22.83333	2.40807	19.40000	27.40000
Unemployment Rate	12	4.72500	0.72504	3.50000	6.00000
Earnings-M	12	35696	4002	29677	41897
Earnings-F	12	25290	2601	20893	29106
Wealth (Income_G_100K)	12	9.43333	2.67525	5.70000	14.40000
Population	12	5366065	4014973	642200	12419293

Methodologies and Factors

In this study, we explore the association of home value with a variety of socio-economic factors, such as, percentage of individuals who attained higher education (bachelor's degree or higher), the unemployment rate, median earnings of male, median earnings of female, percentage of households with income greater than hundred thousand dollars (a measure for wealth) after controlling for the size of the population. Other wealth characteristics, such as, GDP of the state, number of automobiles owned, etc. may be relevant in estimating the housing value. However, they also impact the association of other wealth factors and, therefore, create a confounding effect and thus offset each other in its outcome. For these reasons, they were not considered in this study. Public policy constraints and subsidies that include all types of regulations and taxes also affect the home value through land availability by increasing or decreasing the homeownership incentive. In addition to the land availability or scarcity, one must also consider the influence of regions (location) and the presence of resources that the region can provide. They generate appealing differences between locations and thus create differences in price value. Similarly, mortgage rate that increases or decreases the purchasing power has influence on the price of a home.

TABLE-1C: Summary Statistics of Region3 (South).

Variables	N	Mean	Std Dev	Minimum	Maximum
Home value (Median)	16	96356	22889	70700	146000
BA + %	16	21.35625	4.58330	14.80000	31.40000
Unemployment Rate	16	5.83125	0.90386	4.20000	7.40000
Earnings-M	16	33666	3304	29784	41640
Earnings-F	16	24667	3185	21154	32155
Wealth (Income_G_100K)	16	9.56250	3.72413	5.00000	18.10000
Population	16	6229048	5259687	783600	20851820

In addition, human capital (education) plays an important role on the purchasing power. However, the factor that influences the home value most is the earnings or per capita income. Therefore, the wealth factor (percentage of households with income greater than hundred thousand dollars) is considered as an external factor in our study to observe any wealth dependent effect on the home value. Median male earnings and median female earnings were also analyzed in our study to discern how home value increases with higher per capita wealth. However, the increase of price (value) of a house may be further explained through the effect of other relevant factors. Such as, increase in population size may decrease the land availability and

thus increase the value of the house. Therefore, we control for the population size. Similarly, the unemployment rate was used to control for the differential effect of the health of state's economy. Thus, data on these factors that are stated above were collected for each state from the census bureau and analyzed using associative models. Our research considers modeling the housing relationship of median home value (in dollars) with respect to the wealth effect. In particular, we would like to observe if an increase in home value is associated with the percentage increase in wealthy individuals.

TABLE-1D: Summary Statistics of Region 4 (West).

Variables	N	Mean	Std Dev	Minimum	Maximum
Home value (Median)	13	148869	49531	96600	272700
BA + %	13	24.79231	3.45120	18.20000	32.70000
Unemployment Rate	13	6.21538	1.16107	4.30000	9.00000
Earnings-M	13	36147	3477	30503	41257
Earnings-F	13	26594	3551	20914	31722
Wealth (Income_G_100K)	13	11.33077	3.89987	5.60000	17.30000
Population	13	4861379	8889189	493782	33871648

To observe the association between the home value and the socio-economic factors, initially we observe the bi-variate correlations (results not reported) to examine the direction of the association and the strength of the relationships between factors. Many of these socio-economic factors were statistically significantly correlated with home value. However, they were also highly correlated among themselves due to the fact that they measured the same income (wealth) effect. We then regress home value (median) on the predictors to observe the associations in the housing market. In addition to the inclusion of primary wealth factor, we have also controlled for population size and unemployment rate. As, for example, an increase in population size (more/less) may increase the demand of land and thus result in an increase in home value. Thus, these relevant factors were included during the estimation process of the regression model to control for the proper external effect. In general, it is assumed that there is a difference in home value between good economic condition and difficult economic condition in the process of determining the value of the home and, therefore, the unemployment rate is introduced into the model as an independent variable.

TABLE-1E: Summary Statistics of All Regions.					
Variables	**N**	**Mean**	**Std Dev**	**Minimum**	**Maximum**
Home value (Median)	50	117990	39822	70700	272700
BA + %	50	23.77600	4.28050	14.80000	33.20000
Unemployment Rate	50	5.55400	1.09082	3.50000	9.00000
Earnings-M	50	35830	4252	29677	46368
Earnings-F	50	26091	3473	20893	33318
Wealth (Income_G_100K)	50	10.78800	4.07619	5.00000	21.30000
Population	50	5616997	6185580	493782	33871648

To this end, multiple regression models were run using SAS software (see, SAS/STAT User's Guide, 1993) for housing value on several different factors. These analyses of socio-economic factors are to observe the differential effect on the value of houses due to different regions. This measure is designed to test the hypothesis that home value differences are due to differences in wealth depending on the region.

Specification of the regression model for each region is of the following form:

$$Home\_Value = \beta_0 + \beta_1\, Wealth + \beta_2 Population\ \ Size + \beta_3 Unemployment\ \ Rate + \varepsilon \qquad (1)$$

Specification of the regression model for the whole nation is of the following form:

$$Home\_Value = \beta_0 + \beta_1 Wealth + \beta_2\, \mathrm{Re}\, gion2 + \beta_3\, \mathrm{Re}\, gion3 + \beta_4\, \mathrm{Re}\, gion4 + \beta_5 Population\ \ Size + \beta_6 Unemployment\ \ Rate + \varepsilon \qquad (2)$$

Where:

Home Value: Median home value (in US dollars)

Wealth (Income_G_100K): Percentage of households with income greater than hundred thousand dollars

Population Size: Total population of the state

Unemployment Rate: Unemployment rate of the state

BA+ %: Percentage of Bachelor's degree or higher in the state

Region1: Northeast Region (1, if a state in northeast region, 0 for others)

Region2: Midwest Region (1, if a state in midwest region, 0 for others)

Region3: South Region (1, if a state in south region, 0 for others)

Region4: West Region (1, if a state in west region, 0 for others)

Earnings-M: Median earnings of male residents in the state

Earnings-F: Median earnings of female residents in the state

An increase in either male earnings or female earnings should increase the home affordability and thus increase the home value in that state. Similar effect is expected for an increase in education (human capital). As unemployment rate increases, the total purchasing power decreases and therefore, the home value (in dollars) is expected to decrease. However, these relationships are not independent. To test these hypotheses in our study we have employed associative models in our analysis.

TABLE 2A: Regression results of Home value (median) in Region 1 (Northeast).

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	7007505144	2335835048	9.22	0.0176
Error	5	1266714856	253342971		
Corrected Total	8	8274220000			
R-Square	0.8469		**Adj R-Sq**	0.7551	

Parameter Estimates

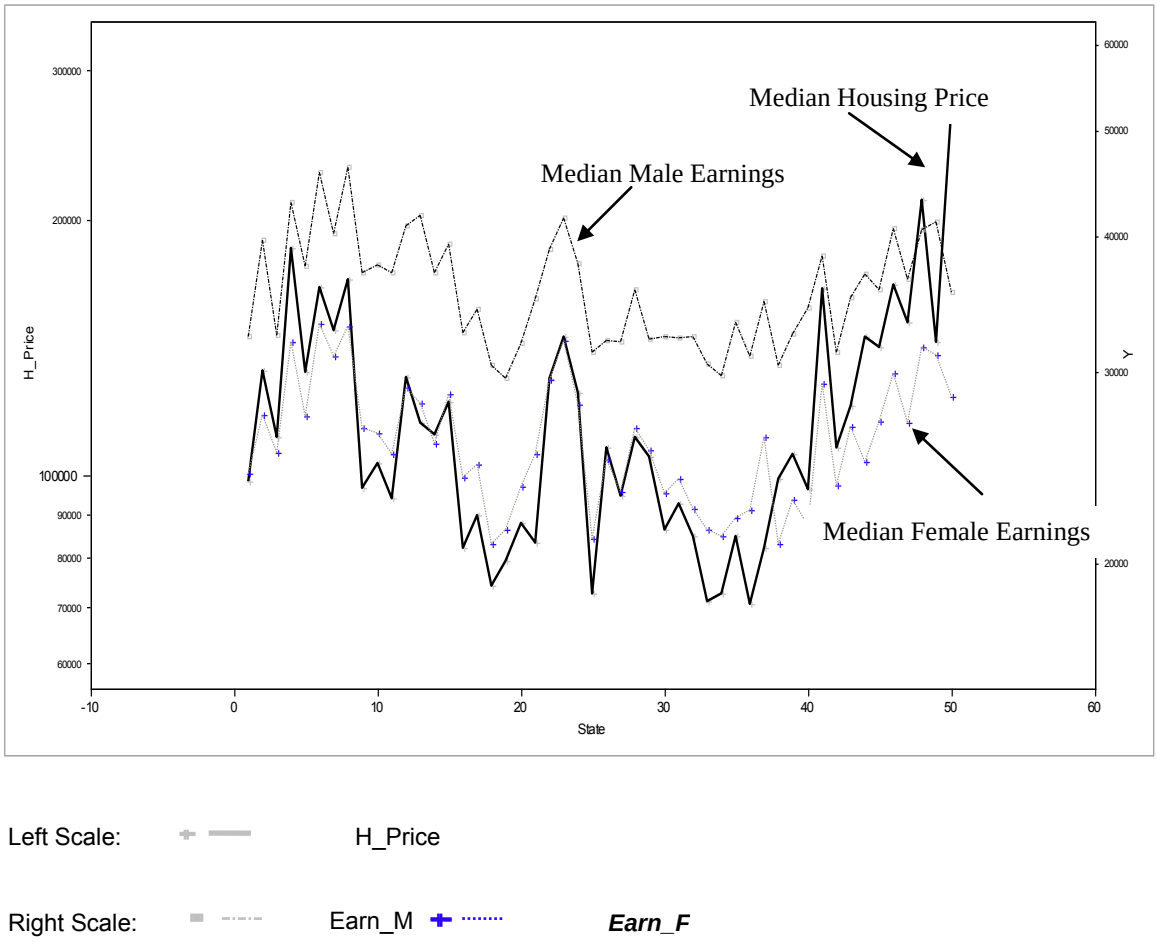
Variables	DF	Parameter Estimates	Standard Error	t Value	Pr > |t|
Intercept	1	65373	45517	1.44	0.2104
Wealth	1	6027.50276	1167.22955	5.16	0.0036
Population Size	1	-0.00022520	0.00154	-0.15	0.8894
Unemployment Rate	1	-1909.40427	9777.58178	-0.20	0.8529

STATISTICAL ANALYSIS

Descriptive statistics for the various measures of factors are calculated and reported in Table-1A through Table-1D for four regions and in Table-1E for the whole nation. West region has the highest average home value ($148,869); whereas south region has the lowest average home value ($96,356) as observed in Table-1D and Table-1C respectively. Relatively larger standard deviations $39,822 (see, Table-1E) of housing value with the highest value being $272,700 in the west region, (see Table-1D) and lowest $70,700 in South region (see Table-1C) do indicate a considerable fluctuation in the home values in different states and thus, indicate much variations in these regions. In addition, percentage of wealth as determined by "percentage of households income greater than 100K" ranges from 5.0 to 21.3 (see Table-1E) reflect substantial purchasing power differences between the states and therefore the regions. In particular, south region has the lowest wealth factor (5.0) and that may be the reason for lowest

home value generated in that region. Given this scenario, we would like to test the hypothesis that the higher the wealth factor (as measured by percentage of household income greater than 100K) the higher the home value. Therefore, we expect a positive relationship between home value and wealth. Higher level of education (percentage of individuals with bachelor's degree or higher) is found to be highly correlated with personal income (results not shown) which also impact housing value. In a similar context, there are also visible similarities and differences in the association between home value vs. median female earnings and median male earnings. In particular, median female earnings are more closely associated with the home value compared to median male earnings as can be seen in Graph-1 and thus exhibiting an important role in home value determination. On the other hand, there is a visible earnings difference (about $10,000) between male and female earners in each region (see, Table-1A-- Table-1D).

Graph-1: Plot of median Home value vs. median Female and median Male Earnings by State.



In addition, simple pair-wise correlation analysis among the variables, reveal that home value is positively impacted by both male earnings and female earnings with a comparatively higher correlation value for female earners. However, the impact on home value is much more pronounced for the wealth factor than any other factors considered. Unemployment rate has opposite and negative impact on the home value though not statistically significant. Average wealth factor is similar in value for the "Northeast" and "West" regions (see Table-1A & Table-1D) and much higher compared to "Midwest" and "South" regions (see Table-1B & Table-1C). Therefore, we expect a negative relationship between the home value and "Midwest" and "South" regions in relation to "West" (see Table-2E), and thus in agreement with our hypothesis of differences in housing value is also due to locational differences that may be dependent on wealth differentials.

TABLE 2B: Regression results of Home value (median) in Region 2 (Midwest).

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	3283647256	1094549085	19.17	0.0005
Error	8	456695244	57086906		
Corrected Total	11	3740342500			
R-Square	0.8779		**Adj R-Sq**	0.8321	

Parameter Estimates

Variables	DF	Parameter Estimates	Standard Error	t Value	Pr > |t|
Intercept	1	48641	24808	1.96	0.0856
Wealth	1	5759.51024	1455.89593	3.96	0.0042
Population Size	1	0.00080989	0.00131	0.62	0.5543
Unemployment Rate	1	-1956.26658	5245.56134	-0.37	0.7189

Results of multiple regression analyses are reported in Table-2A through Table-2E. All these models fit well and are statistically significant in determining the home value. As reported coefficient of determinations (R^2) are 0.85, 0.88, 0.90, and 0.72 for the Northeast, Midwest, South, and West regions respectively, with highly significant F value (see, Table-2A - Table-2D). In addition, an overall model for all four regions combined also produced a statistically significant model with R^2 of 0.81 (see, Table-2E). Results indicate that West region comparative to Midwest or South regions impact the housing value more positively (see, Table-2E). As, for example, home value for a state in the West region is $30,944 higher on average (see, Table-2E), which is also depicted in the summary statistics of home value (see, Table-1D and Table-1C).

However, wealth factor impacts housing value positively across the regions. Analyses also reveal that home value increases by about $6,000 for each additional increase in percentage of households whose income is greater than 100K for three of the four regions, namely Northeast, Midwest, and South regions (see, Table-2A - Table-2C). However, the wealth factor effect for the West region on the home value is almost twice ($11,253) as much compared to other three regions (see, Table-2D). Unemployment rate had a negative effect on the home value; however none of the observed effects of unemployment rate were statistically significant.

TABLE 2C: Regression results of Home value (median) in Region 3 (South).

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	7105020489	2368340163	37.71	<.0001
Error	12	753738886	62811574		
Corrected Total	15	7858759375			
R-Square	0.9041		**Adj R-Sq**	0.8801	

Parameter Estimates

Variables	DF	Parameter Estimates	Standard Error	t Value	Pr > |t|
Intercept	1	61092	25642	2.38	0.0346
Wealth	1	5647.03051	833.22694	6.78	<.0001
Population Size	1	-0.00086164	0.00040722	-2.12	0.0559
Unemployment Rate	1	-2292.47218	3331.20671	-0.69	0.5044

Therefore, in addition to the wealth characteristics, locational differences also affect the housing value differently. Specifically, after controlling for population size and unemployment rate, wealth factor positively impacts the value of a home in each of the regions. Another interesting observation is that the unemployment rate impacted home value differently for different regions (though not statistically significant). As for example, each additional increase in the unemployment rate in the West region impacted the home value negatively three times (see, Table-2D) as much as each of the other three regions. Therefore, after controlling for unemployment rate, this study suggests that the home value is wealth dependent and more importantly the wealth effect is significantly substantial for some of the regions.

TABLE 2D: Regression results of Home value (median) in Region 4 (West).					
Analysis of Variance					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	3	21300694459	7100231486	7.85	0.0070
Error	9	8138733233	904303693		
Corrected Total	12	29439427692			
R-Square	0.7235		**Adj R-Sq**	0.6314	

Parameter Estimates					
Variables	**DF**	**Parameter Estimates**	**Standard Error**	**t Value**	**Pr > |t|**
Intercept	1	63900	49863	1.28	0.2320
Wealth	1	11253	2627.01290	4.28	0.0020
Population Size	1	-0.00008398	0.00112	-0.07	0.9419
Unemployment Rate	1	-6778.74284	7753.28804	-0.87	0.4047

CONCLUSION

In this study, we examine the macro characteristics based analysis that affect the home value to understand the complex nature of housing market dynamics both nationwide and regionally. In particular, statistical significance and magnitude of wealth factor and locational differences on the "home value" is explored. As expected, after controlling for population size and unemployment rate, a higher percentage of wealth is found to be instrumental in affecting the housing value positively and significantly across the regions and, therefore, depicting one of the most important macroeconomic factors' influence on the housing market dynamics. Basic understanding of the relationship of socio-economic factors such as, financial wealth, housing values, and consumption provides a future platform for evaluating economic policy and forecasting economic activity. To this point the observed willingness to purchase higher priced homes by wealthier individuals produces upward price pressure on homes. Changes in house price have a significant impact on gross metropolitan product growth (Miller, Peng, & Sklarz, 2011) as these variables behave differently across different regions. Forecasts about regional trends become possible when similar behavior by the wealthier segment of the population is observed across regions; it may also be possible to infer the performance of the national economy as well. Spending habits of these wealthier individuals are more susceptible to changes in economy and income (Salotti, 2012). For example, wealthy individuals are more likely to pay premium prices for homes with socially desirable property names during stronger economic conditions than weak economic conditions (Zahirovic-Herbert & Chatterjee, 2011). Future

research should examine the wealth effects carefully as high net worth households may behave differently than high income individuals that may have a lower percentage of net worth tied to financial assets (Guo & Hardin, 2014). Overall, a small portion of the home buyer population in these findings proved to be instrumental in influencing home prices, and future research should identify its operational ability on the economic impact.

TABLE 2E: Regression results of Home value (median) in All Regions.

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	63030930912	10505155152	30.78	<.0001
Error	43	14674374088	341264514		
Corrected Total	49	77705305000			
R-Square	0.8112		**Adj R-Sq**	0.7848	

Parameter Estimates

Variables	DF	Parameter Estimates	Standard Error	t Value	Pr > |t|
Intercept	1	45484	19828	2.29	0.0267
Wealth	1	7164.61674	784.96754	9.13	<.0001
Region2 (Midwest)	1	-8367.66898	9022.43587	-0.93	0.3589
Region3 (South)	1	-9456.45318	8572.80316	-1.10	0.2761
Region4 (West)	1	30944	8728.44855	3.55	0.0010
Population Size	1	0.00000953	0.00048800	0.02	0.9845
Unemployment Rate	1	-1413.43847	2964.80005	-0.48	0.6360

Thus, these results add another dimension in this field of research concerning the importance of the wealth factor on the home value that has a regional differential effect. In addition, the magnitude of the wealth factor's impact on the home value is location dependent. Specifically, after adjusted for population size and unemployment rate, wealth factor impacts the value of the house positively across the regions. This particular association between wealth and housing value is an important finding of this paper. Although, the data indicate much variability in the home values for different regions, effect is substantially positive even after adjusted for socio-economic factors. However, the extent of the wealth effect is different for different regions. Thus, this study provides evidence for policy makers and regulators to understand some of the underlying forces in the housing market movement and may be valuable for future policy making process.

REFERENCES

Algieri, B. (2013). House price determinants: Fundamentals and underlying factors. *Comparative Economic Studies, 55,* 315-341.

Block, J.P., R.A. Scribner, & K.B. DeSalvo (2004). Fast food, race/ethnicity, and income: A geographic analysis. *American Journal of Preventive Medicine, 27*(3), 211-217.

Boelhouwer, P.J., M.E.A. Haffner, P. Neuteboom, & P. de Vries (2001). House prices and income tax in the Netherlands: An international perspective. *Housing Studies, 19*(3), 415-432.

Campbell, A.L. & M.W. Sances (2013). State fiscal policy during the great recession: Budgetary impacts and policy responses. *The Annals of the American Academy of Political and Social Science, 650*(1), 252-273.

Campbell, J.Y. & J.F. Cocco (2007). How do house prices affect consumption? Evidence from micro data. *Journal of Monetary Economics, 54*(3), 591-621.

Case, K. E., J. M. Quigley, & R.J. Shiller (2005). Comparing wealth effects: The stock market versus the housing market. *Journal of Macroeconomics: Advances in Macroeconomics, 5*(1), 1-32.

Clark, D.E. & W.E. Herrin (2000). The impact of public school attributes on home sale prices in California. *Growth and Change, 31*, 385–407.

Davidoff, T. (2003). Labor income, housing prices and homeownership. *Journal of Urban Economics, 59*(2), 209-235.

Downes, T.A. & J.E. Zabel (2002). The impact of school characteristics on house prices: Chicago 1987–1991. *Journal of Urban Economics, 52*, 1–25.

Duda, M., & E. Belsky (2002). Asset appreciation, timing of purchases and sales, and returns to low-income homeownership. In: Retsinas N., Belsky, E. (eds) *Low Income Homeownership: Examining the Unexamined Goal* (pp. 208-238). Brookings Institution, Washington.

Feldman, R. (2002). The affordable housing shortage: Considering the problem, causes and solution. *Banking and Policy,* Working Paper 02-2, Federal Reserve Bank of Minneapolis.

Figlio, D. & M. Lucas (2004). What's in an output? School report cards and the housing market. *American Economic Review, 94*(3), 591–604.

Fullerton Jr, T.M. & E. Villalobos (2011). Street widths, international ports of entry and border region housing values. *Journal of Economic Issues, 45*, 493-509.

Fullerton, T. M., K.P. Fierro, & E. Donjuan-Callejo (2009). Housing attribute preferences in a northern Mexico metropolitan economy. *Atlantic Economic Journal 37*, 159-172.

Garcia, J., D. Montolio & J.M. Raya (2010). Local public expenditures and housing prices. *Urban Studies, 47*(7), 1501-1512.

Guo, S. & W.G. Harding III (2014). Wealth, composition, housing, income and consumption. *Journal of Real Estate Finance and Economics, 48*(2), 221-243.

Holmes, M. J., J. Otero & T. Panagiotidis (2011). Investigating regional house price convergence in the United States: Evidence from a pair-wise approach. *Economic Modelling, 28*(6), 2369-2376.

Kenny, L.W. & A. Reinke (2011). The role of income in the formation of new cities. *Public Choice, 149*, 75-88.

LaVeist, T. & J. Wallace (2000). Health risk and inequitable distribution of liquor stores in African American neighborhood. *Social Science Medicine, 51*, 613-617.

Malpezzi, S. (2003). Hedonic pricing models: A selective and applied review. In T. O. Sullivan & K.Gibbs (Eds.), *Housing Economics and public policy: Essays in honor of Duncan Maclennan.* Oxford, UK: Blackwell.

Miller, N., L. Peng, & M. Sklarz (2011). House prices and economic growth. *Journal of Real Estate Finance and Economics, 42*(4), 522-541.

Morland K., S. Wing, A. Diez Roux, & C. Poole (2002). Neighborhood characteristics associated with the location of food stores and food service places. *American Journal of Preventive Medicine 22*, 23-29.

Poterbal, J.M. & T. Sinai (2011). Revenue costs and incentive effects of the mortgage interest deduction for owner-occupied housing. *National Tax Journal, 64*(2), 531-564.

Reback, R. (2005). House prices and the provision of local public services: Capitalization under school choice programs. *Journal of Urban Economics, 57*(2), 275–301.

Riley, S.F. (2012). Land use regulations and the returns to low-income homeownership. *Annals of Regional Science, 49*, 745-766.

Rosenthal, S.S. (2008). Old homes, externalities, and poor neighborhoods: A model of urban decline and renewal. *Journal of Urban Economics, 63*, 816–840.

Rouwendal, J. & S. Longhi (2008). The effect of consumers' expectations in a booming housing market: Space-time patters in the Netherlands, 1999-2000. *Housing Studies, 23*(2), 291-317.

Salotti, S. (2012). Wealth effects in the US: Evidence from the combination of two surveys. *Economic and Social Review, 43*(1), 67-98.

Seo, Y. & R. Simons (2009). The effects of school quality on residential sales prices. *Journal of Real Estate Research, 31*(3), 307–327.

Sirmans, S.G. & D.A. Macpherson (2003). The state of affordable housing. *Journal of Real Estate Literature, 11*(2), 131-156.

Sirmans, G.S., D.A. Macpherson, & E.N. Zietz (2005). The composition of hedonic pricing models. *Journal of Real Estate Literature, 13*(1), 3-43.

Turner, T.M. & M.T. Smith (2009). Exits from homeownership: The effects of race, ethnicity, and income. *Journal of Regional Science, 49*(1), 1-32.

Zahirovic-Herbert, V. & S. Chatterjee (2011). What is the value of a name? Conspicuous consumption and house prices. *Journal of Real Estate Research, 33*(1), 105-125.

Zhu, B., R. Füss, & N.B. Rottke (2013). Spatial linkages in returns and volatilities among U.S. regional housing markets. *Real Estate Economics, 41*(1), 29-64.

Zietz, J., E.N. Zietz, & G.S. Sirmans (2008). Determinants of house prices: A quartile regression approach. *Journal of Real Estate Finance Economics, 37*, 317-333.