Volume 14, Number 3

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., 51 Blake Drive, Arden, NC 28704, 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

LETTER FROM THE EDITOR

The *JEEER*, the official journal of the Academy of Economics and Economic Education, is dedicated to the study, research and dissemination of information pertinent to the discipline of economics, and to the improvement of methodologies and effective teaching in economics. The *Journal* bridges the gap between the theoretical discipline of economics and applied excellence relative to the teaching arts. The Academy is an affiliate of the Allied Academies, Inc., a non profit association of scholars whose purpose is to encourage and support the advancement and exchange of knowledge, understanding and teaching throughout the world.

The Editorial Board considers two types of manuscripts: first is theoretical and empirical research related to the discipline of economics. The second area is research oriented toward effective teaching methods and technologies in economics.

These manuscripts have been double blind reviewed by the Editorial Board members. The manuscripts published in this issue conform to our acceptance policy, and represent an acceptance rate of 25% or less.

We are inviting papers for future editions of the *Journal* and encourage you to submit your manuscripts according to the guidelines found on the Allied Academies webpage at www.alliedacademies.org.

Grady Perdue
University of Houston-Clear Lake

Martin Milkman
Murray State University

John Marcis
Coastal Carolina Univeristy

EVALUATING THE EFFECTS OF TAXING THE REMITTANCES OF SKILLED WORKERS ON CAPITAL ACCUMULATION AND AGGREGATE INCOME USING AN OVERLAPPING GENERATIONS MODEL[1]

John Paolo R. Rivera, De La Salle University

ABSTRACT

Labor migration has sizable economic impacts specifically to labor-sending countries such as the Philippines, which can alter the economy's production structure and redirect the country's comparative advantage. The exodus of highly trained professionals, without replacement, will lead to brain drain in a country with limited access to quality higher education especially if the education costs of these professionals have been subsidized by the state; hence, a substantial loss to society is incurred. Likewise, the training costs of replacements can be reasonably substantial and may cause the reduction of the productivity of workers left behind. Thus, this study developed an Overlapping Generations (OLG) Model on the Philippine context that will discuss the management of skilled labor migration and assessing the macroeconomic effects it entails. By hypothetically incorporating how a tax on the income of skilled migrant workers abroad specifically to those schooled in state universities and colleges (SUCs), as proposed by Bhagwati (1976), affects the macroeconomy, this study provides an insight on the efficacy of its implementation. Simulation results have shown that imposing the brain drain tax can enable the economy to achieve a higher steady state capital stock and steady state aggregate income paths on the condition that the government will not spend all the revenues from the brain drain tax on one generation.

INTRODUCTION

It has been established that labor migration has non-negligible economic impacts specifically to labor-sending countries such as the Philippines. It is because the labor-sending economy can be subjected to the incidence of brain drain and experience the ills and benefits of remittance flows. As explained by Mandelman & Zlate (2009), temporary labor migration varies over the business cycle due to the prevalence of cyclical unemployment brought about by booms and busts in both the labor-sending and labor-receiving economies. Mandelman & Zlate (2009) provided evidence that there are drastic declines in labor immigration flows during recession in most developed countries, which was supported by the findings of Tullao, Conchada & Rivera

(2010) wherein there is evidence of changing demand for nursing graduates and other professional workers in the Philippines due to the global crisis that occurred for the past decade.

With more than 10 percent of the population stationed as either permanent residents, or temporary workers, or illegal migrants in more than 182 countries, the Philippines has emerged as one of the major exporters of labor services in the world (Collymore, 2003). According to Sjaastad (1962), labor migration is a household investment decision that depends on the incentive to migrate which is reliant on the expectations of future earnings at the destination country relative to the country of origin. As such, due to the lack of job opportunities and the unattractive compensation packages in the Philippine labor market, many have opted to seek employment abroad in the hopes of augmenting household domestic income. As the country continues to struggle with political and economic instabilitues, the continued exodus of Filipino labor, especially skilled labor, will continue to prevail.

According to the Philippine Overseas Employment Agency (POEA), there have been 1,470,826 deployed Overseas Filipino Workers (OFWs[2]) across the globe in 2010 with a 3.4 percent increase from 2009. It is vital to note that it is difficult to measure with any precision the exact number of Filipinos working abroad. Figures from the POEA only include those people who are working abroad with registered and official contracts, while the number of Filipinos working abroad on an irregular and unofficial basis is unknown but probably quite high. Nonetheless, these figures are likely to persist if there will be no major change within the country's economic policies.

As a consequence of temporary labor migration, the labor-sending country is able to receive remittance income. According to Adam & Page (2005), McKenzie & Sasin (2007), and Acosta, Lartey & Mandelman (2009), the magnitude and the growth rate of remittances received by various developing economies has exceeded the inflow of official aid and foreign direct investments (FDIs). In recent years, that the value of remittances in 2005 was approximately 2.5 percent of gross national income (GNI) in the developing world (Acosta, Lartey & Mandelman, 2009). Subsequently, according to the World Bank (2006) as cited by Acosta, Lartey & Mandelman (2009), the large magnitude of remittance income has contributed to the reduction of absolute poverty, the improvement of human capital indicators, and the reduction of income inequality. However, according to Stark (1988) and Barham & Boucher (1995), migration and remittances has worsened income inequality as compared to a no-migration counterfactual. Moreover, according to Tuaño-Amador, Claveria, Delloro & Co (2008), remittances are counter-cyclical in nature that unlike FDIs and capital inflows, during times of crisis, the amount of remittance still increases. This is because migrant workers must send financial support for the survival of their families. Hence, this protects the economy from further recessions since remittances provide for the consumption expenditure of the recipient country.

Meanwhile, in the Philippines, the OFWs send remittances to their respective households in the country on a regular basis that stimulated the economic performance of the Philippines despite the occurrence of the financial crisis (Tullao & Rivera, 2008). Moreover, according to

Tullao, Cortez & See (2007), the remittances that augments investment in human capital show that migration has both social benefits and social costs. For social benefits, the remittance inflow from OFWs recorded by the Bangko Sentral ng Pilipinas (BSP) in 2010 at USD 18.762 billion contributed on the subsequent improvement in economic and social status of families with members who are working abroad and are often cited as major positive contributions of international labor migration. Likewise, on the average, remittances account for more than 12 percent of the country's Gross Domestic Product (GDP), which signals that economic growth in the Philippines is externally induced (Ang, Sugiyarto & Jha, 2009). Similarly, according to Mandelman & Zlate (2009), the insurance role of remittances in smoothing the consumption path of receiving households is also a contributory factor in their welfare enhancement. Hence, it is evident that remittance has become the principal component of total household financial inflows.

On the other hand, the incidence of labor migration in the Philippines also poses social costs, which on the other hand is mitigated by the way higher education in the Philippines is financed in the more than 1,400 institutions of higher learning. According to Tullao, Cortez & See (2007), almost 80 percent of the students in higher education are attending private educational institutions. Since the cost of higher education is privately financed, demand for higher education can be seen as an investment to enhance their chances of migration. This is the concept of the culture of migration discussed by Tullao & Rivera (2008) wherein because of the success of their family members in global employment, the other members of the family particularly the young ones may also want to seek external employment. Since in the global labor market, the preferred and highly paid workers are the more educated than the less educated ones, there is a tendency for families to invest in education as a means of increasing the chances of their family members to seek overseas employment. In addition, they see future remittances as private returns. Accordingly, one way of compensating the country for the loss of migrants, who attended government funded state universities and colleges (SUCs), is to internalize the cost of their education. Another option is to impose some form of exit tax on migrating workers like nurses whose massive exit have affected nursing education as well as the health sector of the country (Tullao, Cortez & See, 2007).

Then again, it may be argued that aside from the monetary costs, the opportunity costs of overseas employment may also be quite substantial. Massive migration can alter the structure of production in the sending communities and redirect the country's comparative advantage. If the laborers leaving the country are the skilled ones, the training costs of replacements may be quite substantial and may cause the reduction of the productivity of workers left behind. The exodus of highly trained professionals, without replacement, will lead to brain drain in a country with limited access to quality higher education. If the costs of education of these professionals have been shouldered by the state, a substantial loss to society is incurred when these professionals migrate permanently (Tullao & Cortez, 2003).

Indeed, labor migration is such a huge phenomenon that affects the various areas of living. Various studies such as that of Tullao, Cortez & See (2007), Tullao & Cabuay (2011), and

Ducanes (2011) have already established the economic impacts of overseas migration on Philippine households and the aggregate economy. Likewise, Ducanes (2011) also highlighted that overseas labor migration can affect households, macroeconomy, society, human rights, and other political facets of the sending country. The immensity of this phenomenon is forcing the government to implement controlling procedures to impede the exodus of manpower that will arrest the possible hollowing effects on industries and mitigate the loss in international competition as argued by Tullao, Cortez & See (2007). Hence, this study will focus on the labor migration of skilled professionals, which can result to the incidence of brain drain. Specifically, this study is interested in determining the distortions and/or the welfare-enhancing effects of the imposition of a brain drain tax on the emigration of skilled labor as proposed by Bhagwati (1976) and by Tullao, Cortez & See (2007). This study is running on the premise that the possibility of increasing and internalizing the cost of international migration can reduce the economic ills it has generated. In particular, revenues from these initiatives to impose taxes on emigrating skilled workers can be channeled to improve the productivity of workers left behind. As such, this study has the following specific research objectives:

- To develop a theoretical overlapping generations (OLG) model that will incorporate how a hypothetical tax on the income of skilled migrant workers abroad schooled in SUCs affects the paths of capital accumulation and aggregate income of the economy; and
- To assess the theoretical OLG model using calibration in order to determine the implications of the tax reform. This methodology will allow the determination of the effect of taxes on capital accumulation and aggregate income.

In the light of the emerging culture of migration among Filipino households, it becomes an important research inquiry to determine whether the initiatives to control brain drain are plausible and whether it is able to minimize economic costs and harness the benefits of this important contemporary phenomenon on the sending country. Likewise, this study becomes even more significant as the Philippines is one of the labor exporting countries in the world and has relied on substantial remittances from OFWs for sustaining stability and growth of the economy.

MANAGING SKILLED LABOR MIGRATION THROUGH TAXATION

International skilled labor migration is inevitable as evidenced by the alarming figures that ought to trigger authorities to manage and regulate the flow of labor through imposition of rules. This should not be limited to the sending countries only but also to receiving countries as well. According to Lowell, Findley & Stewart (2004), there three alternative areas wherein a governing body can generate policies that will revolve on managing labor migration. These are migration management, the "diaspora option" and democracy and development.

Likewise, the incidence of labor migration through remittances has always been perceived as a primary driver of economic growth and development (Tullao & Rivera, 2008). It does not only enhance consumption in the microeconomic level, but it also serves as a medium to uplift the quality of life of remittance-dependent households. Yet, the distinguishable benefits through the high volume of remittances are tantamount to the accompanying social costs of migration. The government ceases to see the impact of the negative externalities of migration since there is no government policy yet that will manage labor migration specifically the exodus of skilled labor.

As defined by Lowell, Findley & Stewart (2004), migration management focuses on the temporary movement of human capital by establishing regulations in receiving countries through admission policies and sending countries in the form of efficient practices and creating an environment that will attract the return of migrants. It control out-migration from at-risk countries, accountability for recruitment agencies and employers, best practice for employing foreign workers, temporary worker schemes and facilitate and create incentives for return.

Migration management policies are specific policies geared towards mitigating the societal costs of migration such as the incidence of brain drain, the Dutch Disease phenomenon, and the distortions in the labor market. Several known migration management policies proposed include the levying of a brain drain tax, the provision of schooling incentives, and imposition of work bonds to those who benefited from the government's human resource development programs and educational subsidies.

As presented by Mirrlees (1971), taxing migrants is an effective way to distribute income from high skilled workers to low skilled workers. Hence, the study came into the conclusion, that migration of high skilled workers creates "less egalitarian tax system". Increasing tax burden on skilled residents is one of the considerations why they migrate which results to loss of revenues. Egger & Radulescu (2009) modeled how high tax rates should be imposed to highly skilled residents before considering or deciding to migrate. It showed that the most important component is the "progressivity of tax system" in high income brackets. Individuals in this bracket are mostly the highly skilled workers. Then, in the implementation of taxing skilled migrants, it is suggested that to know how elastic migration is with respect to changes in taxes in home countries (Wilson, n.d.). The idea of taxing migrants, aside from the compensation due to loss of revenues, is another factor to be considered by residents whether to leave or not since the presence of tax is another cost for them. Hence, the brain drain tax is not just another redistribution process but a constraint in the flow of migration. According to Scalera (2006, 2009), the Bhagwati Tax, after considering externalities and a functioning government, is a welfare-improving system because the tax may actually boost investment on human capital.

The literature has cited the various social costs when skilled workers migrate from home country. Through taxation the government of the sending country can compensate from the loss of skill. Since instead of benefiting from the possible contributions of the home-trained skilled

worker, other countries are profiting and the home country is left behind. There has been a revolutionary taxation scheme proposed by Bhagwati (1976).

Bhagwati (1976) proposed to impose an exit tax, also known as the Bhagwati tax, upon the departure of skilled labor migrants from their home country. Its basic principle is to tax migrants from the loss of skilled manpower. It is suggested that this type of tax is to be collected "under UN auspices". Then the United Nations (UN) will be the one responsible in the allocation of tax revenues and will not consider the "corrupt and dictatorial" countries. The revenue allocation will be based on the developmental programs proposed by the sending countries (Bhagwati & Dellafar, 1973).

This type of collection and tax imposition is crucial and difficult to implement. Also, it sparked numerous criticisms because it was deemed unfair and unjust. Bhagwati (1979) then revised his proposition and reported that tax collection should be done by the developing countries or the sending countries through the use of global tax. Furthermore, Bhagwati (1979) emphasized that taxation and its benefits must go to the sending country because migrants are able to retain their nationality and rights. Hence, without appropriates taxation relating to migrant mobility, there is "representation without taxation".

Most sending countries are developing countries and the implementation of such sophisticated tax system might be another channel of ineffective income redistribution. An on-going argument is that, non-benevolent government in developing countries might use the brain-drain tax as another collection that will simply end in non-profit earning investments. Furthermore, taxing citizens of the source country while in the receiving countries might end in double taxation cases which will be deemed more harmful to migrants. This procedure might pose a great challenge to governments, especially governments of developing countries, seeing that tax collection in their own home countries is proving to be difficult as well. It is then suggested by Wilson (2003) that a fixed flat rate should be charged to citizens abroad in order to answer the practicality issue of tax collection.

As such, according Wilson (2003), despite the considerable amount of literature supporting the idea of taxing emigrants, the major roadblock in the implementation of this tax is due to administrative issues specifically for the developing nations. Since problems would exists when taxes are levied on foreign-source income. Bhagwati (1979) urged that developing nations could collect a tax from skilled emigrants as a form of a global tax system, wherein foreign and domestic incomes are both taxed. However, only the United States (US) and a few countries attempted to tax foreign income and they had a hard time given that they have a highly developed tax system. One of the main reasons with the difficulty of taxing foreign income is that there must be information sharing among governments (Wilson, 2003). Wilson (2003) also proposed probable solutions in administering the tax through a voluntary brain drain tax. The tax system was proposed so that returning emigrants face lower tax payments once they return to their home country particularly if they have previously paid the brain drain tax, as compared to those who evaded the brain drain tax. He found out that emigrants would voluntarily pay the

brain drain tax rather than risk the probability of paying taxes once they return home; non-emigrants should also be taxed which in overall represents a residence-based tax on skilled labor. The main idea behind this is that emigrating does not affect a skilled worker's lifetime tax burden. Ideally, everyone should faced the same brain drain tax but he realized that income of skilled workers vary abroad, thus creating a role of brain drain tax which varies across emigrants. One probable solution is for the emigrant to supply their income abroad in their tax reforms which can be used by the government to implement varying brain drain tax. Also, if the income abroad is positively correlated with domestic income, the government can implement varying tax penalties as a form of brain drain tax to returning emigrants.

On the other hand, Scalera (2009) reviewed the initial idea of the Bhagwati tax and argued that when taking into account the social externalities of migration, the implementation of Bhagwati tax tends to foster human capital formation and increase resident's income and welfare. Also, if brain drain tax is paid other than the normal income tax, fiscal burden can be outweighed by higher human capital and gross income. Scalera (2009) also discussed the possibility that the tax can easily be avoided because of the initial proposal that the tax collected by the receiving countries shall route revenues towards sending countries. However, taxing poor migrants can be seemingly odious and non-discriminatory which is why receiving countries are less likely willing to levy and transfer a brain drain tax. If brain drain occurs (i.e. more skilled workers migrant than unskilled workers), the mean value of human capital decreases therefore reducing society's overall welfare and income per capita. In such case, a Bhagwati tax can be used to internalize migration benefits and induce the social planner to aim for a higher optimal human capital. Knowing that there are a lot arguments against the Bhagwati tax, Scalera (2009), did not criticize the proposal of Bhagwati but instead argued when taking into account social externalities of human capital and government policies caring for only those residents who are left behind, brain drain tax can be beneficial to all agents; wherein a fiscal burden could be outweighed by higher human capital and income per capita while destination countries can also find it profitable to channel a part of the migrant's income abroad to sending countries in exchange of higher skilled immigration.

Meanwhile, Brauner (2010) examined the levying of a brain drain tax as a development policy regime. It examined the potential of taxation in generating development funds in accordance with international skilled migration from developed to developing countries. Brauner (2010) discussed that the implementation of brain drain is considered to be impossible to administer but with the present international tax regime, brain drain taxation may actually be administered. A study by Kapur & McHale (n.d.) as cited by Brauner (2010) argued that the Bhagwati tax must be re-examined due to changes in circumstance such as the accessibility on the data of the magnitude of brain drain. Likewise, the changes in administrative environment, globalization, trans-nationalism, and international cooperation are sufficient reasons to warrant another attempt in implementing the tax. Moreover, citizenship-based taxation is much more feasible since citizenship is "worth more" and information furnishing from developed countries

are not prohibited. Furthermore, Kapur & McHale (n.d.) as cited by Brauner (2010) provided further support for the brain drain tax by analyzing the whole spectrum of potential policy implementations pertaining to brain drain. Findings included a citizenship-based taxation, which followed the modern version of the Bhagwati tax - a revenue sharing model which requires cooperation among developed countries and an exit tax mechanism.

The most attractive forms of tax is a citizenship-based taxation and a taxation system which involves cooperation amongst receiving and sending countries (Brauner, 2010). It is still apparent for the former that developing countries might not have the ability to administer the tax and the latter cannot be materialized without actual cooperation with the developed nations. Bhagwati (1976) proposed the imposition of surtax, a tax levied on income, implemented by the host country but it was deemed impossible or rather difficult since there is unequal taxation of these immigrants as compared to those residents in the home country. The next alternative is a tax imposed by the host country. However, it was problematic since tax jurisdiction follows residence based on our current international tax regime. Fortunately, a solution for this taxation was found in the form of an alternative personal jurisdiction regime, which is the citizenship based taxation. The main problem for this resolution is the imposition of this citizenship based worldwide income tax is constrained by legal and political complications (Brauner, 2010).

Meanwhile, this does not mean that the tax is infeasible to impose by the sending country. It can still be valid if the sending country currently imposes a tax on a worldwide basis. If the legal system of countries consider immigrants as their residents once they have entered the territory, then the taxation is possible. Moreover, it is also feasible if there are established tax treaties, which are open for minor amendments to incorporate the proposed Bhagwati style brain drain tax. The tax treaties envision cases of dual residence wherein immigrants, by laws, remain to be residents of their home countries and include a provision to "break the tie" such that the immigrant is only considered to be a resident of one of the countries involved (Brauner, 2010). Generally, taxpayers have the capability of altering their dealings in order to produce a positive benefit from taxation and it is likely for skilled migrants to avoid normal residence taxation by their own respective countries. This is due to the fact that income taxation in developing countries that are usually the labor sending countries is progressive (Todaro & Smith, 2006). A probable approach to this is to implement a separate tax outside the scope of the treaty but this approach faces the same problem of citizenship based taxation. Another possible approach is to amend the "tie breaking rules" which in return considers immigrants resident of their own respective host countries to avoid now the incidence of double taxation (Brauner, 2010).

In conclusion, the success of a tax implementation can be measured by its effectiveness in promoting development. Brauner (2010) highlighted that existing literature only focused on the taxing mechanism of brain drain tax and not how the collected revenue can be put into work. Bhagwati (1976) avoided a much more straightforward solution wherein taxes are directly transferred to sending countries or mixing them with host countries' foreign aid funds. Bhagwati (1976) argued that tax proceeds must be thought about more carefully since blending them with

general funds utilized by the international organizations will be counterproductive, and the "new" tax may be viewed as another excuse to increase aid or yet contribute to the increasingly wasteful list of aid mechanisms. However, Bhagwati (1976) suggested a much more simpler solution through a bilateral agreement, wherein host countries would divert the brain drain tax revenues to sending countries and the former would comply solely because for the purposes of supporting development. The revenue generated must not be combined with the general tax budget since it ceases to see direct effectiveness of the newly generated tax. Conversely, countries may choose to formulate a bilateral vehicle which would review the use of tax proceeds and audit its consistency with the rules and regulations established in the agreement. Lastly, Brauner (2010) suggested other tax mechanisms in the hope of implementing the brain drain tax namely exit taxes and tax sharing. Exit taxes serve as a deterrent for brain drain since it makes labor migration much more costly that spurred objections since it restricts movement of individuals in moral and human rights grounds. Also, this tax mechanism is ineffective since emigrants are yet to benefit from the increase in wages constituting to a positive benefit in the long run. On the other hand, revenue sharing is a tax mechanism that must be explored further because it requires sophisticated cooperation between the host and sending countries regarding their respective tax policies and enforcement levels. The present reality of non-tax cooperation, even among developed countries makes it skeptical if this policy is even viable.

THE OVERLAPPING GENERATIONS MODEL

The OLG model developed by Diamond (1965) will be utilized following the standard assumptions. In this model, one young and one old generation exists at any point in time. Assume that individuals in this two-period economy work full time when young and are retired when old. Moreover, on the first period, an individual must decide whether he or she will work in the domestic labor market or abroad. In period 2, all those who worked abroad referred to as the migrant will return to the home country and retire. Also assume that neither the population nor productivity grows and, for the moment, that there is no government. The young choose their current consumption and anticipated old age consumption on the basis of their preferences and their lifetime resources. Likewise, the young during period 1, by altruism, contributes to the consumption of their relatives. Since parents in this life cycle model are assumed to spend their old age resources, which are comprised of their savings, income earned on their savings, and out of the money given by their offspring who are in their first period working full time, entirely on their old age consumption, there are no bequests, gifts, or other forms of net intergenerational transfers to the young. As a result, the young have no nonhuman wealth, and the lifetime resources of the young correspond to the labor earnings they receive when young.

This study adopts the convention that output is produced, income is received , and consumption occurs at the end of each period, the tangible wealth of the economy at the beginning of any period consists of private assets held by the elderly. Since the elderly consumes

all available resources in their possession at the end of their last period of life, the capital stock available to the economy in the next period consists of savings by the current young that they bring into the next period, which is their old age.

Thus, the supplies of productive factors to the economy consist of the labor supply of the current young plus the capital supplied by the elderly, which is the savings of last period's young generation. These factors are supplied to the production sector of the economy. The output of the production sector in turn is paid out to the productive factors as returns to capital and labor. In the model, equity and debt are perfect substitutes. Hence, the elderly are completely indifferent between exchanging their capital for stocks or bonds at the beginning of their second period and receiving a return of principal plus capital income in the form of dividends and proceeds from the sale of their shares in the case of stocks and in the form of interest plus principal payments in the case of bonds. Furthermore, since the production sector is assumed to be competitive, factors of production specifically labor and capital are hired to the point where marginal revenue products equal factor payments. For the economy to be in equilibrium, the time path of factor demands must equal the time path of factor supplies.

The Components of the Model

Consider a two-period model in which the Philippine economy is represented by a utility function and domestic production function given by Equation 1 and Equation 2.

$$U_t^A = \ln c_t^A c_t^B + \left(\frac{1}{1+\delta}\right)\ln\left(c_{t+1}^A - \gamma c_t^A\right) \tag{1}$$

$$Y_{dt} = K_{dt}^{\alpha} L_{dt}^{1-\alpha} \tag{2}$$

Equation 1 expresses the inter-temporal utility of a member of generation t, U_t^A, as a function of his or her monetary-valued consumption when young, c_t^A; monetary-valued consumption of his or her relatives during the first period, c_t^B, by virtue of altruism; $1/(1+\delta)$ as the subjective discount factor; and δ, which is the pure rate of time preferences. Auerbach & Kotlikoff (1975) defined the pure rate of time preference as the degree to which, other things being equal, the individual would prefer leisure and consumption in an earlier rather than later year. It is also vital to note from Equation 1 that in the presence of habit formation, the utility of a given level of consumption when old must not be independent of consumption when young. According to Wendner (2000), the absolute level of consumption in the second period as well as the increase of second period consumption relative to first period consumption are important. Wendner (2000) furthered that the more that was consumed when young, the more is required to derive the same level of utility in the following period, which is technically referred to as habit

formation or habit persistence. As such, habit formation is represented by $c_{t+1}^{A} - \gamma c_{t}^{A}$ where γ denotes the strength of habit formation. This formulation of habit formation is quite standard in the literature as exemplified by Wendner (2000).

On the other hand, the Philippine economy's production function, as seen in Equation 2, relates output per young worker, Y_{dt}, to capital per young worker, K_t, and labor per young worker, L_t. It is assumed that L_t is exogenously supplied by each young worker and is measured in units such that $L_t = 1$.

Equation 3 gives the inter-temporal budget constraint of an individual who is young at time *t*.

$$\left(c_t^A + c_t^B\right) + \frac{c_{t+1}^A - \gamma c_t^A}{1+r_d} = \left[m_{dt}^A + (1-\tau)m_{ft}^A\right] + \left[\frac{\rho_1 m_{d(t+1)}^B + \rho_2(1-\tau)m_{f(t+1)}^B}{1+r_d}\right] \tag{3}$$

Since this study is interested in the imposition of a brain drain tax on skilled Filipino migrant workers, specifically to those who graduated from SUCs, as recommended by Bhagwati (1976), the inclusion of this fiscal policy alters the model in two ways. First, it changes lifetime budget constraints, with after-tax lifetime resources substituted for their pre-tax values. Second, the capital stock now corresponds to total national net wealth, that is, the net wealth of the government plus the private sector.

Assuming the role of the social planner, suppose that the brain drain tax system is designed in such a way that the skilled labor migrants will have to pay a constant proportional tax on his or her income earned abroad for all *t*. Assume that a specific tax τ, where $1 > \tau > 0$, is imposed on m_{ft} earned by skilled migrant labor. As long as the skilled Filipino migrant labor is working abroad, the brain drain tax will be applied on his or her wages to be collected by the host foreign economy and shall be remitted to the sending country via an existing facility agreed by both economies. Further assume that a certain proportion of the tax revenue is returned to the people in the form of government transfers, and other welfare-enhancing services.

Note that $\left(m_{dt}^A + m_{ft}^A\right)$ is the income that a young member of generation *t* earns at the first time period where m_{dt}^A is the income earned by working in the domestic market and m_{ft}^A is the income earned by working abroad, which will be used to finance c_t^A and c_t^B. If the individual decides to work in the domestic labor market, then $m_{dt}^A > 0$ and $m_{ft}^A = 0$. On the other hand, if the individual decides to work abroad, then $m_{dt}^A = 0$ and $m_{ft}^A > 0$.

On the second period, the consumption of the individual is now financed by the endowment provided by the younger generation, who are currently on their first period, by altruism. If the younger generation decides to work in the domestic labor market, then $m_{d(t+1)}^B > 0$

and $m^{B}_{f(t+1)} = 0$. On the other hand, if the younger generation decides to work abroad, then $m^{B}_{d(t+1)} = 0$ and $m^{B}_{f(t+1)} > 0$. The parameters ρ_1 and ρ_2 denotes the proportion of income earned by the younger generations in the second period allocated for the consumption of the older generation. All values in the second period are discounted by the domestic interest rate, r_d. Assuming perfect capital mobility, domestic interest rate must equal foreign interest rate.

Also, it can be deemed that a portion of m^{A}_{ft} and $\rho_2 m^{B}_{f(t+1)}$ are the remittances sent to the home country in the first period and second period respectively. See Tullao & Cabuay (2011) for the details on the motivations for sending remittances. With the role of remittances, it is now evident that the role of migration is incorporated in the optimization problem. As argued by McKenzie & Sasin (2007), one actually cannot and in most cases separate remittances from migration, because these phenomena are intertwined and endogenous. In fact, it is not immediately clear why one would want to separate them and what the pure impact of remittances would mean or imply. Likewise, the constraint assumes a linear form since working abroad and working in the home country are deemed to be perfect substitutes.

Implementing optimization procedures with the abovementioned equations will yield consumption demands, savings function, and capital accumulation path. Because of the mathematical rigors involved in this study, only the resulting equations are presented. As such, Equation 4 shows the law of motion of capital.

$$K_{d(t+1)} = \frac{\left(1+\alpha K_{dt}^{\alpha-1}\right)(1-\tau)(1-\beta)K_{ft}^{\beta} - 2(1+\delta)\left[\rho_2(1-\tau)(1-\beta)K_{f(t+1)}^{\beta}\right]}{(3+2\delta)\left(1+\alpha K_{dt}^{\alpha-1}\right)} + \Lambda_{dt}^{g}\left(1+\alpha K_{dt}^{\alpha-1}\right) + \tau n(1-\beta)K_{ft}^{\beta} - \theta Y_{d(t-1)} \tag{4}$$

It is important to note that the government's choice of the time path of its consumption and tax instruments is constrained by its inter-temporal budget constraint. This constraint requires that the present value of the government's outlay equals the present value of its receipts plus its initial net worth. While restricting the set of feasible policies, the government's long term budget is consistent with a wide range of short- and medium-term policies. In particular, the government can permit debt to grow for a long time at a faster rate than the economy, although indefinite use of this policy is not feasible in this model, since under such a policy debt would eventually exceed national wealth and the capital stock will be negative.

For any particular government policy, the perfect foresight assumption requires that household correctly foresee the time path of government policy variables entering their budget constraints. That is, the generation that is young at time *t* must correctly foresee both r_d and τ. Perfect foresight, although not required by the model, is needed by the government if it is to implement effectively its desired fiscal program.

Suppose the economy is initially in a steady state. The equation for the economy's initial steady state capital stock is shown by Equation 5, which can be solved for in an essentially non-algebraic way. Note that under the steady state, $K_{d(t+1)} = K_{dt} = K$ and the rest are parameter values.

$$K = \frac{\left(1+\alpha K^{\alpha-1}\right)(1-\tau)(1-\beta)K_{ft}^{\beta} - 2(1+\delta)\left[\rho_2(1-\tau)(1-\beta)K_{f(t+1)}^{\beta}\right]}{(3+2\delta)\left(1+\alpha K^{\alpha-1}\right)} + \Lambda_{dt}^{g}\left(1+\alpha K^{\alpha-1}\right) + \tau(1-\beta)K_{ft}^{\beta} - \theta Y \tag{5}$$

Setting $K_{d(t+1)} = K_{dt} = K$ will yield a non-linear equation in the steady state capital stock; hence, the solution to this equation may not be unique and has no closed form solution.

Assuming that the young migrates on the first period and his or her offspring also migrates on the second period by culture of migration among Filipino households; Equation 5 can be collapsed into Equation 6.

$$K = \frac{\left(1+\alpha K^{\alpha-1}\right)(1-\tau)(1-\beta)K_{ft}^{\beta} - 2(1+\delta)\left[\rho_2(1-\tau)(1-\beta)K_{f(t+1)}^{\beta}\right]}{(3+2\delta)\left(1+\alpha K^{\alpha-1}\right)} \tag{6}$$

Suppose that in this two-period life cycle model, the Philippine economy is initially in a steady state in which there is no government debt and government consumption is financed by an income tax from the households working in the domestic economy. Moreover, assume that the young migrates on the first period and his or her offspring also migrates on the second period by culture of migration among Filipino households. Thus, the solution for the economy's initial steady state capital stock can be obtained from Equation 6 where τ is the steady state income tax rate under the assumption that Λ_{dc}^{g} equals zero.

Suppose the government in this economy, as per advised by the social planner, announces a ten-period imposition of a brain drain income tax, that is, the government will impose a fixed tax rate for the next 10 periods and that the tax revenues are to be spent on government transfers on welfare-enhancing services. After the 10th period, another tax policy regime will be implemented wherein the government will impose a higher fixed tax rate for the subsequent periods.

If the brain drain tax is announced at time $t = 0$, the equation for the economy's capital stock at $t = 1$ is given by Equation 7

$$K_1 = \frac{\left(1+\alpha K^{\alpha-1}\right)(1-\tau)(1-\beta)K_{ft}^{\beta} - 2(1+\delta)\left[\rho_2(1-\tau)(1-\beta)K_{f(t+1)}^{\beta}\right]}{(3+2\delta)\left(1+\alpha K^{\alpha-1}\right)} + \Lambda_1^{g} \tag{7}$$

where Λ_1^g is the government's net assets at time t =1. Meanwhile, the formula for Λ_1^g is given by Equation 8

$$\Lambda_1^g = \tau Y_f - \theta Y_d \tag{8}$$

where Y_f is the initial steady state level of income in the host economy where the brain drain tax revenues will be sourced, Y_d is the initial steady state level of income in the domestic economy, and θ is the proportion of steady state income that the government spends on welfare-enhancing projects. For $t \geq 2$, the economy's capital stock is determined by Equation 5 with $\Lambda_{dc}^g = \Lambda_1^g$.

The general equilibrium economic model this study utilizes forms the basis for all the simulation results to be generated. This section examines the choice of parameter values and the method of solving for the quantities and prices that characterize the perfect foresight equilibrium this study is following.

The Simulation Method

The calculation of the equilibrium path of the economy, given a particular parameterization, typically proceeds in three stages as enumerated by Auerbach & Kotlikoff (1987) namely solving for the long run steady state of the economy before the assumed change in fiscal policy begins; solving for the long run steady state to which the economy converges after the policy takes effect; and solving for the transition path that the economy takes between these two steady states.

The perfect foresight assumption is important only in the third stage, since in either of the long run steady states economic variables are constant from one year to the next; any plausible assumption about the formation of expectations would lead individuals to have correct foresight in such situations. The transition begins when information about the policy change becomes available. One should visualize this as an unanticipated change in the fiscal policy regime (Auerbach & Kotlikoff, 1987).

Following the methodology of Auerbach & Kotlikoff (1987), households and firms have perfect foresight in both old and new policy regimes, but do not anticipate the policy change. The policy change may take the form of immediate changes in fiscal variables or of immediate announcements of future changes in fiscal variables. In the case of pre-announced policies, the transition also begins in year 1, which is the year always used to index the beginning of the transition. Although there is no change in fiscal policy until several years later; that is, households and firms has perfect foresight about the future switch in regime, in pre-announced policy changes, the transition begins as soon as the future policy is announced.

Simulation analysis is the only alternative available when it is necessary to analyze large policy changes in models that are too complicated for simple analytical solutions. To solve the OLG model, values for the different parameters identified and used in this study must be chosen. If the model is to be as realistic as possible, the numerical estimates of these parameters must be culled from the empirical literature and some parameters would require an indirect method to obtain their values. What is important is that all assumed parameter values have basis.

According to Auerbach & Kotlikoff (1987), simulating the model for alternative policies replaces the comparative static procedures that are performed with analytical models. Moreover, one can implement sensitivity analysis of the numerical simulation model by examining the impact of conceivable variation in parameter values. Oftentimes, the results of such sensitivity analysis are robust to parameter changes, even though this outcome cannot be foreseen prior to performing the simulation procedure. In other cases, results are quite sensitive to small changes in particular parameters, which is deemed by Auerbach & Kotlikoff (1987) as useful information because it is indicative of which parameters must be precisely estimated empirically.

Table 1: Parameter Values for Calibration

Parameters	Description	Value	Source
A	domestic capital's share of domestic output	0.1627	Derived using the average share of Gross Domestic Capital Formation to Gross National Income from 2001 to 2010 sourced from the Bangko Sentral ng Pilipinas. World Bank figures revealed an average of 0.1840 from 2006 to 2010.
β	foreign capital's share of foreign output	0.2033	According to Auerbach & Kotlikoff (1975), it is well known for a Cobb-Douglas production function that factor shares are constant, and the capital share in income equals the capital-intensity. Using the historical share of capital in national income, from 2006 to 2010 sourced from the World Bank, for the countries where skilled Filipino labor are deployed namely Saudi Arabia, United Arab Emirates (UAE), Hong Kong, Italy, United Kingdom (UK), Canada, and the United States of America (USA).
ρ_2	proportion foreign income earned by skilled migrant workers to be remitted to the domestic economy	0.7	It is assumed that 70 percent of the incomes of skilled migrant workers are remitted to their households in the Philippines.
δ	pure rate of time preferences	0.7353	There is a limited evidence of the appropriate value of δ. As such, this study derived its value using the average saving rate of the Philippines from 2001 to 2010 sourced from the BSP. Upon determining the average saving rate, the marginal propensity for current consumption was determined to represent the pure rate of time preferences. This value was selected because it leads to a realistic consumption profile and labor supply decision for reasonable tax parameters and levels of government consumption in the Philippines.
γ	strength of habit formation	0.4	Otrok (1999)

Table 1: Parameter Values for Calibration			
Parameters	**Description**	**Value**	**Source**
K_{ft}	foreign capital per young worker at time t	1	The steady state foreign capital was normalized to 1 on the assumption that developed countries where OFWs are deployed are more developed than the Philippines; hence, they would have a higher level of steady state capital stock than the Philippines.
$K_{f(t+1)}$	foreign capital per young worker at time t+1	1	
τ	brain drain tax as a proportion of income earned abroad by the skilled migrant worker	0.05, 0.10	The proper choice of tax bases is a central question in tax reform. The choice has important implications for the course of savings and economic growth, the distribution of welfare across generations, and the level of economic efficiency in the economy. Hence, a sensitivity analysis on the steady state capital will be implemented under these values of τ. Likewise, the actual tax system is characterized by a tax base that is much narrower than national income, and hence by much higher marginal tax rates than would be suggested by these revenue percentages. The effect of changing the tax base is one of the issues explored in the simulation procedures to be implemented.
θ	proportion of steady state output that the government spends	0.05, 0.06	A sensitivity analysis on the steady state capital will be implemented under these values of θ. In the Philippine setting, based on the historical data of the government spending and GDP from the NSO, government spending absorbs a varying 5 to 15 percent over a period of 2 decades. These choices of values represent a compromise made necessary by the simplicity of the model relative to the real economy. Likewise, these values appear quite reasonable in terms of overall revenue.

Table 1 shows the assumed values for the parameters utilized in this study, specifically for calibrating the model with the role of government, together with the basis for choosing the values. Recall that the parameters were declared in the utility function and aggregate production function shown by Equation 1 and Equation 2 respectively.

Given the parameterization of the OLG model designed by this study, an exact numerical solution for the equilibrium of the economy for any given fiscal policy can be obtained, which can be compared with the results for different fiscal policies. By itself, this is the essence of the numerical simulation approach.

To show the movement of the steady state consumption, capital, and income if the government will impose a brain drain tax on skilled migrant workers, calibration procedures will have to be implemented again wherein all parameters in Equation 5, taken as exogenous, will assume the values shown in Table 1.

To analyze the dynamic fiscal policy transition, suppose that in this two-period life cycle model, the Philippine economy is initially in a steady state in which there is no government debt and government consumption is financed by an income tax from the households working in the domestic economy. Note that the economy starts at t=0. Suppose also that the government is not yet imposing a brain drain tax. Further suppose that at the initial period, the government spends 5 percent of the initial steady state output.

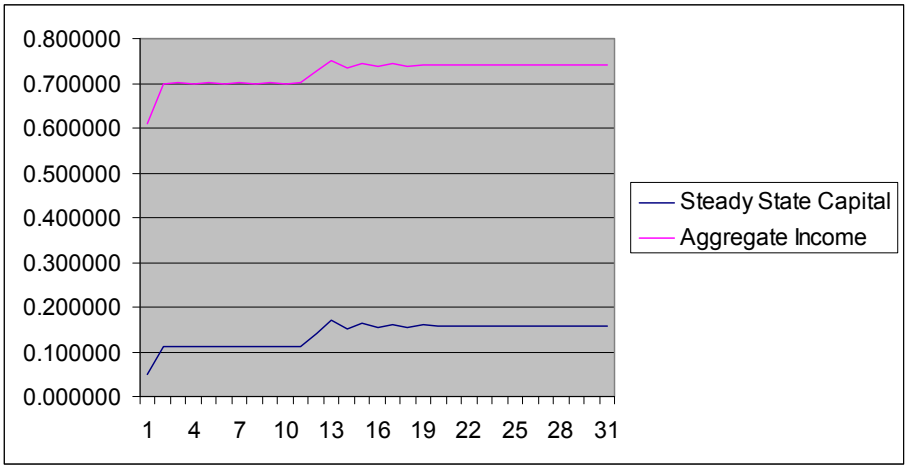
Suppose the Philippine government, as per advised by the social planner, announces a ten-period imposition of a brain drain income tax, that is, the government will impose a fixed tax rate of 10 percent for the next 10 periods, starting at t=1 to t=10, on the income of skilled migrant workers abroad, specifically to those schooled in SUCs, and that the tax revenues are to be spent on government transfers on welfare-enhancing services, which is assumed to be 5 percent of the initial steady state output.

After the tenth period, which is at t=11, another fiscal policy regime will be implemented wherein the government will impose a higher fixed tax rate for the subsequent periods. This study will look into a specific state of the world to analyze the consequences of imposing the brain drain tax and further increasing the tax rate on steady state capital and steady state aggregate income; that is, the government increases the brain drain tax rate from 10 percent to 15 percent and it will also increase its proportion of steady state output spending from 5 percent to 6 percent.

The Resulting Steady State Paths of Capital Accumulation and Aggregate Income

After implementing calibration procedures, Figure 1 graphs the impact of the brain drain tax on the paths of steady state capital stock and steady state output values for the next 30 periods. From the apparent values found in Figure 1, the number may seem strange to those accustomed to thinking of capital-output ratios as being between 3 and 6 as typified by Auerbach & Kotlikoff (1987), wherein the numbers make sense once one realizes that a period in this two-period model corresponds to roughly 30 years in real time. Therefore, output in a two-period model is roughly equivalent to output over a 30-year period, and one must multiply the two-period capital-output ratio by 30 to arrive at a roughly equivalent annual figure.

Figure 1: Impact of the Brain Drain Tax Policy on Steady State Capital and Aggregate Income



It can be observed from the results that the imposition of the brain drain tax on a macroeconomic level caused capital stock and aggregate income to increase in order to reach a higher level of steady state value until a new fiscal policy is implemented. Moreover, implementing a new tax policy by t=11 will further allow the economy to reach a much higher steady state capital stock and steady state income with some bumps after the new tax policy is imposed before it reaches the new and higher steady state. This is evidently the transition path that the economy takes between these two steady states. It can be interpreted as the period in which workers adjust consumption due to the tax as well as the period in which the economy adjusts to the policy shock.

CONCLUSIONS

The Philippines has one of the best practices of temporary labor migration in the globe as evidenced by the volume of OFWs, with varying skills ranging from semi-skilled to highly skilled, being deployed in various destinations and the flow of remittances, which are all considered as the major agents of economic stabilization. Herewith, labor migration undeniably has extensive and significant economic impacts especially to labor-sending economies. In the midst of temporary labor migration, any labor-sending economy can experience brain drain, which is operationally defined as the exodus of skilled labor without them being replaced by the domestic economy. The incidence of brain drain has been found by various literatures to alter the economy's production structure and redirect the country's comparative advantage. Likewise, the migration of highly trained professionals, devoid of replacement, will lead to brain drain in a country with limited access to quality higher education especially if the education costs of these professionals have been subsidized by the government such as those from various SUCs; hence, a substantial loss to society is acquired. Likewise, the training costs of replacements can be reasonably substantial and may cause the reduction of the productivity of workers left behind.

Consequently, this study explored the inter-generational effects of levying a brain drain tax on capital accumulation and aggregate output by developing an OLG model specifically on the Philippine context. The foundation of this study's model is from a simple two-period life cycle model with an inter-temporal utility function of the skilled migrant worker that simultaneously incorporates altruism and habit formation. The production sector, the inclusion of the government, and all other assumptions followed the standard framework of an OLG model.

Simulation results have shown that imposing the brain drain tax, and increasing it after an extended period of time, can enable the economy to achieve a higher steady state capital stock and steady state aggregate income paths on the condition that the government will not spend all the revenues from the brain drain tax on one generation. Hence, the results of calibrating and simulating the OLG model developed engendered the various classic results from an OLG framework such as the convergence to a new steady state path, the divergence from the steady

state path, and the trade-offs in welfare between and across generations. Accordingly, for social planners, the imposition of a brain drain tax and spending a portion of it on the provision of public goods can allow the macroeconomy to reach a higher steady state capital accumulation and steady state aggregate income for an extended period of time.

One may argue that there is no point being belligerent about the issue of brain drain tax especially under the idea that temporary migration brings about temporary brain drain. Conversely, under the OLG framework, it must be noted that regardless of whether the skilled labor returns or not and retires by the second life period, that individual will eventually die, which can be deemed as permanent brain drain, and the economy will not have any use for that individual. Also, the fact that the economy was not able to fully harness the productivity of that skilled labor; it warrants the need for a compensation policy for the loss to society. However, the necessity for the brain drain tax is subject to a lot of considerations that the social planner must deliberate specifically the obvious decline in consumption as a result of an additional taxation. Hence, the predicament of the social planner is between the welfare of the households or macroeconomic growth and the welfare of the current generation or the succeeding generations.

ENDNOTES

[1] This study was culled from the author's Doctoral Dissertation, for the degree Doctor of Philosophy in Economics in De La Salle University, entitled *The International Migration of the Highly Skilled Filipino Labor: A Theoretical Consideration of the Welfare and Macroeconomic Impacts of Taxes on Remittances*.

[2] According to Tullao, Cortez & See (2007), migrant worker refers to a person who is to be engaged, is engaged or has been engaged in a remunerated activity in a state of which he or she is not a legal resident; to be used interchangeably with Overseas Filipino Worker per Republic Act (RA) 8042 also known as the Migrant Workers and Overseas Filipinos Act of 1995.

REFERENCES

Acosta, P.A., E.K.K. Lartey & F.S. Mandelman. (2009). Remittances and the Dutch disease. *Journal of International Economics*, 79(2009), 102-116.

Adams, R.H. & J. Page. (2005). Do international migration and remittances reduce poverty in developing countries? *World Development*, 33, 1645-1669.

Ang, A., G. Sugiyarto & S. Jha. (2009). *Remittances and household behaviour in the Philippines*. ADB Economics Working Paper Series 188.

Auerbach, A.J. & L.J. Kotlikoff. (1987). *Dynamic fiscal policy.* Cambridge University Press.

Barham, B. & S. Boucher. (1998). Migration, remittances, and inequality: Estimating the net effects of migration on income distribution. *Journal of Development Economics,* 55, 307-331.

Bhagwati, J.N. (1976a). *Taxing the brain drain, volume 1: A proposal.* Amsterdam: North-Holland.

Bhagwati, J.N. (1976b). *The brain drain and taxation, volume 2: Theory and empirical analysis.* Amsterdam, North-Holland.

Bhagwati, J.N. (1979). International migration of the highly skilled: Economics, ethics and taxes. *Third World Quarterly*, 1, 17-30.

Bhagwati, J.N. & W. Dellafar. (1973). The brain drain and income taxation. *World Development,* 1, 94-100.

Brauner, Y. (2010). Brain drain taxation as development. *Social Science Research Network.* Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1636845

Collymore, E. (2003). *Rapid population growth, crowded cities present challenges in the Philippines*. Retrieved from http://www.prb.org/Articles/2003/RapidPopulationGrowthCrowdedCitiesPresentChallengesin the Philippines.aspx

Diamond, P.A. (1965). National debt in a neoclassical growth model. *American Economic Review,* 55, 1126–1150.

Ducanes, G.M. (2011). *The welfare impact of overseas migration on Philippine households: A critical review of past evidence and an analysis of new evidence from panel data*. Doctoral Dissertation. University of the Philippines – Diliman.

Egger, P. & D. Radulescu. (2009). *The influence of labor taxes on the migration of skilled workers*. Retrieved from http://www.cesifo-group.de/portal/page/portal/DocBase_Content/WP/WP-CESifo_Working_Papers/wp-cesifo-2008/wp-cesifo-2008-11/cesifo1_wp2462.pdf

Lowell, L., A.M. Findley & E. Stewart. (2004). *Brain strain - Optimizing highly skilled migration from developing countries.* Institute for Public Policy Research.

Mandelman, F.S. & A. Zlate. (2009). *Immigration and the macroeconomy: An international business cycle model.* Retrieved from the Proceedings of the BSP International Research Conference on Remittances.

McKenzie, D. & M.J. Sasin. (2007). *Migration, remittances, poverty and human capital: Conceptual and empirical challenges.* World Bank Policy Research Working Paper 4272.

Mirrlees, J.A. (1971). An exploration in the theory of optimal income taxation. *Review of Economic Studies,* 38, 175-208.

Otrok,C. (1999). *On measuring the welfare cost of business cycles.* University of Virginia.

Scalera, D. (2006). *Education, policies, brain drain, and Bhagwati tax.* Retrieved from http://economia.uniparthenope.it/ise/sito/conferenze/miurVinci/Papers_Accepted/Scalera_bdtax.pdf

Scalera, D. (2009). *Skilled migration and education policies: Is there still scope for a Bhagwati tax?* Università del Sannio, Department PEMEIS, Piazzetta Arechi II, 82100 Benevento, Italy

Sjaastad, L.A. (1962). The costs and returns of human migration. *The Journal of Political Economy*, 70(5), 80-93.

Stark, O. (1988). Migration, remittances, and inequality: A sensitivity analysis using the extended Gini index. *Journal of Development Economics*, 28, 309-322.

Todaro, M. P. & S.C. Smith. (2006). *Economic development 8th edition*. Philippines: Pearson South Asia Pte. Ltd.

Tuaño-Amador, M., R. Claveria, V. Delloro & F. Co. (2008). *Philippine overseas workers and migrants remittances: The Dutch disease phenomenon and cyclicality issue.* Bangko Sentral Ng Pilipinas.

Tullao, T.S. (2008). *Epekto ng padalang salapi ng mga OFW sa tunay na palitan ng salapi*. Proceedings of the Fourth Ambassador Alfonso Yuchengco Policy Conference. Yuchengco Center. De La Salle University.

Tullao, T.S. & C.J.R. Cabuay. (2011). *International migration and remittances: A review of economic impacts, issues and challenges from the sending country's perspective.* Angelo King Institute.

Tullao, T.S. & J.P.R. Rivera. (2008). *The impact of temporary labor migration on the demand for education: Implications on the human resource development in the Philippines*. East Asian Development Network

Tullao, T.S & M.A.A. Cortez. (2003). *The movement of natural persons between the Philippines and Japan: Issues and prospects*. Philippine APEC Study Center Network. Philippine Institute for Development Studies

Tullao, T.S., M.A.A. Cortez & E.C. See. (2007). *The economic impacts of international migration: A case study on the Philippines.* East Asian Development Network.

Tullao, T.S., M.I.P Conchada & J.P.R. Rivera. (2010). *The labor migration industry for health and educational services: Regulatory and governance structures and implications for national development.* World Bank

Wendner, R. (2000). *A policy lesson from an overlapping generations model of habit persistence.*Austrian Science Foundation. Manuscript.

Wilson, J.D. (n.d.). *Taxing the brain drain: A reassessment of the Bhagwati proposal*. Michigan State University. Retrieved from http://bear.warrington.ufl.edu/dinopoulos/Bhagwati/PDF/Wilson-R.pdf.

Wilson, J.D. (2003). A voluntary brain drain tax. *Journal of Public Economics*, 92(12), 2385-2391.

THE ONE VILLAGE ONE PRODUCT (OVOP) MODEL AND ECONOMIC DEVELOPMENT ON GUAM

Ning Li, University of Guam
Fred R. Schumann, University of Guam

ABSTRACT

Guam is a small island state in the western Pacific Ocean. The island's tourism industry has served as an engine for its economic growth. However, the local community of Guam is not maximizing the benefits of tourists' spending due to a high level of leakage. This paper proposes that the One-Village-One-Product (OVOP) strategy be implemented on Guam so that Guam residents may benefit more from the tourism industry by building up linkages with goods and services suppliers on Guam. Candidate products for each village on Guam are identified, indicating that Guam is ready for a One-Village-One-Product strategy. Policy recommendations are given for fostering Guam's tourism industry.

INTRODUCTION

Guam is a tropical island in the western Pacific Ocean. It is the southernmost and largest island in the Mariana island chain. It is also the largest island in Micronesia. Yet Guam is still considered a small island state measured by its geographical size and population. Guam has an area of 212 square miles (549 km2). Inhabiting this tropical island are approximately 160,000 people with diverse racial and cultural backgrounds. In terms of racial breakdown of Guam, the largest group is Chamorro, which accounts for 37.1% of the total population. Other groups are Filipino (25.5%), White (10%), Chinese, Japanese and Korean (8%), and Pacific Islanders and others (20%).

Besides commerce between the island's inhabitants, like many other small islands around the world, the tourism industry serves as one of the major sector of the island's economy. Guam receives more than one million tourists annually, among which about 80% are from Japan, about 10% are from Korea, and the rest are from Taiwan, Hong Kong, United States, and etc.

Unlike many of the small island states, Guam's economy also benefits from its military-related service sector. Due to its strategic location, Guam hosts two military bases of the United States in this island, namely the Andersen Air Force Base and the U.S. Navy Joint Region Marianas. In addition, it has been forecasted that the impending relocation of U.S. Marines from military bases in Okinawa, Japan, to Guam may bring drastic changes to this island's socio-economic landscape. However, the changing political, social, and economic dynamics in Japan

and the United States complicate this military relocation and, as a result, many critical issues relative to this military buildup remain unclear.

This paper posits that although the proposed military buildup may provide good opportunities for Guam to boost its economy and enhance the living standard of Guam people, the extent and timing of the military relocation to Guam is still fraught with high level of uncertainty. On the other hand, Guam's tourism industry provides a more promising path of Guam's long-term economic growth. Further, this paper argues that implementing the One-Village-One-Product (OVOP) strategy will bring about positive results for Guam's economic growth by bringing in more tourist arrivals and encouraging tourist spending with minimal leakage of revenue. The remaining parts of this paper will review the significance of Guam's tourism industry, investigate whether Guam is ready for an OVOP strategy, and make some policy recommendations for fostering Guam's tourist industry.

DEVELOPMENT ECONOMICS FOR SMALL ISLAND ECONOMY

Economic growth and structural change are key concerns of development economics. Various theories and methods that provide guidance for public policy-making and implementation at both domestic and international levels have been created. The aim is to ensure economic growth for developing economies and to improve living standard of the population. Among these theories, two strands of literature in development economics deserve special mention: the five-stage growth model and the structural change theory.

The five-stage growth model was proposed by W.W. Rostow. In his classics *The Stages of Economic Growth: A Non-Communist Manifesto*, Rostow (1960) asserts that all countries must go through a series of five consecutive stages of development: (1) the traditional society in which products are primarily consumed by producers; (2) the preconditions for take-off under which increased specialization generates surpluses for trading; (3) the take-off, when industrialization increases with concentration in one or two manufacturing industries; (4) the drive to maturity, when the economy is diversifying into new areas; and (5) the age of high mass-consumption, when service sector dominates the economy. According to Rostow, the primary means of promoting economic growth is through accelerated accumulation of capital. In other words, development requires substantial investment in capital. Rostow's model may be good for most nations of the world, but it does not always apply to small island states such as Guam. Most small island states do not have enough savings and thus lack necessary capital goods capacity to drive economic growth.

The structural change theory focuses on the change of economic structures in the process of the transition from a primarily agricultural society into a modern, urbanized, industrial, and service economy. Unlike Rostow's assertion that all nations follow the same path of growth, the structural change theory holds that different nations become wealthy through different trajectories. Chenery (1960, 1968) argues that a nation's choice of trajectories of growth

depends on various factors such as size, resource endowment, population, current income level, and comparative advantages relative to other nations. Demas (1965), in his study of Caribbean nations, concludes that very small countries should follow a different pattern of growth from that of large continental countries. He argues that, since it is generally difficult for small nations to ensure capital goods capacity and internally self-sustaining growth, these nations should seek externally oriented growth engines.

Among small countries, small island states manifest some unique characteristics. Various studies have identified typical advantages and disadvantages of small island economies. Here, small island economies are defined as island states with less than one million inhabitants and less than 2,000 square miles (or 5,000 km^2) in area. There are close to fifty small island economies in the world.

On the negative side, size constraints are the major considerations. First, because of their small domestic markets, economies of scale are typically difficult to obtain (Knox, 1967). Second, the geographic location which often specifies remoteness and isolation brings about many difficulties for development in small islands (Brookfield, 1990). This is still true today, even though technological advances have produced state-of-the-art air transportation systems. Third, due to their susceptibility to natural disasters (typhoon, earthquake, tsunamis, etc.), small island states suffer from macroeconomic instability, especially if the island has a strategy of export specialization and concentration (Briguglio, 1995). This is why small island states are often characterized with import dependence. Finally, in terms of social and cultural impacts, some (Armstrong and Read, 1998; Benedict, 1967) argue that, due to intense kinship ties and ethnic heritages, policy consensus is often difficult to achieve and decision-making is often rather subjective than objective.

On the positive side, small island states generally endowed with natural beauty such as sun, sea, sand, and cliff (Aguilo et al, 2005). Island demography with the propensity to migrate and progress is another positive perspective, as it will in turn bring back remittances and financial aid (Bertram and Watters, 1985), and it will provide long-run metropolitan links (Bertram, 2004). In addition, due to their small size, these island states can take advantage of relatively easily implementing domestic policy in restructuring island economies towards service sectors (i.e., tourism and off-shore banking) and towards export-oriented manufacturing (McElroy and Mahoney, 2000). An empirical study by McElroy (2006) proves promoting tourism industry is a viable strategy for economic growth in small island states.

SIGNIFICANCE OF THE TOURISM INDUSTRY ON GUAM

Guam's tourism industry has served as an engine for its economic growth. Tourism has produced jobs and income for people, generated revenue and profits for business, and harvested tax money for government. In addition, tourism has brought in investment and even business relations and new residents to Guam, as it has been known to do in many destinations (Kotler et

al, 2002). The tourism industry has helped Guam obtain a self-sustaining economy and higher standard of living.

Guam still lacks statistics and calculations on its macroeconomic indicators. The recent estimate of Guam's gross island product (GIP) was around $4 billion (Ruane, 2012). The Government of Guam is considered a major employer on the island. The two critical sectors that contribute largely to Guam's economy are U.S. military and tourism. Military-related businesses as well as those that indirectly provide goods and services are expected to benefit with the planned relocation of U.S. Marines from Okinawa, Japan to Guam. However, it is still unclear when and how many marines will be relocated. On the other hand, the tourism industry is more dependable and sustainable as a source of revenue for the island. The average number of tourist arrivals in recent years has been around 1.2 million. The ratio of tourist arrivals to inhabitants on Guam is about 7.5, higher than most major tourist destinations.

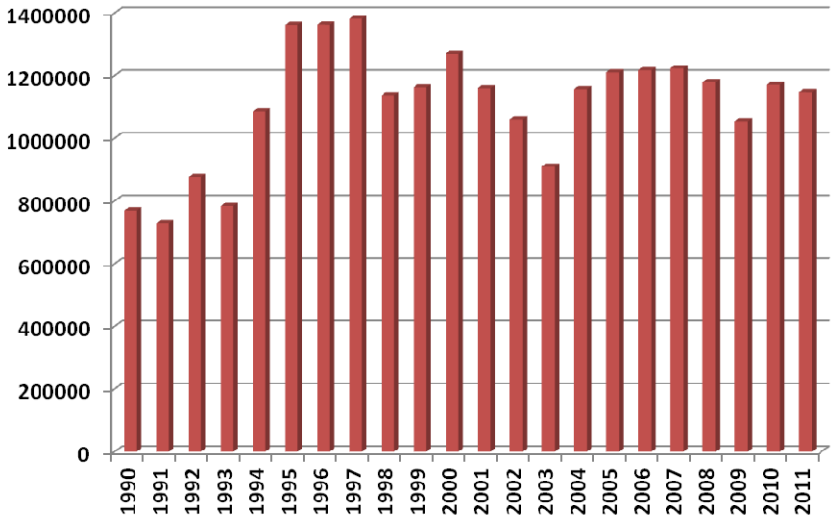
Guam's tourism industry began to take off in the 1960s with the growth of air travel. The first commercial airplane loaded with international tourists to Guam was from Japan operated by Pan American Airways in the year 1967. Other airlines followed soon after. In late 1980s and early 1990s, the tourism industry on Guam benefited tremendously from the buildup of hotels in the area of Tumon Bay, with most of the investments coming from Japan. Today, Guam is the home of many famous international hotels chains, including Nikko, Westin, Marriott, Hilton, Sheraton, Hyatt, and Outrigger brands. Figure 1 shows tourist arrivals over the years from 1990 to 2011.

The tourism industry on Guam is vulnerable to natural disasters, socio-economic situations in the source markets, and, to some extent, the international social, political, and economic performance. Figure 1 reflects those impacts. For example, the Gulf War in 1991, the terrorist attack on September 11 of 2001, the Iraq War in 2003, and SARS outbreak in 2003-04 imposed negative influences to Guam's tourist arrivals. The significant appreciation of Japanese Yen against U.S. dollars in mid-1990s brought in a significant increase of Japanese tourist to Guam. In 1997, Guam reached its peak in tourist arrivals with the record of 1.38 million. The Asian financial crisis in 1997-1998, joined by Typhoon Paka in December 1997, dragged tourist arrivals down by almost 300,000. Natural disasters also brought about negative impacts. Examples included Typhoon Omar in August 1992, an 8.2 Earthquake in August 1993, Typhoon Chataan, Halong, and Ponsonga in 2002. A more recent event with negative impact was the 2011 Earthquake in Japan.

For small island economies, the tourism industry has the potential to contribute significantly to economic growth and increase income. According to a study conducted by Global Insight, Inc. (Pike, 2007), Guam's tourism industry accounts for 20 percent to one-third of the island's aggregated economy. In the year 2005, the tourism industry generated 20% of Guam's GIP or 32% of Guam's non-government GIP, paid 36% of non-government wages, and supported 35.5% of the private employment base.

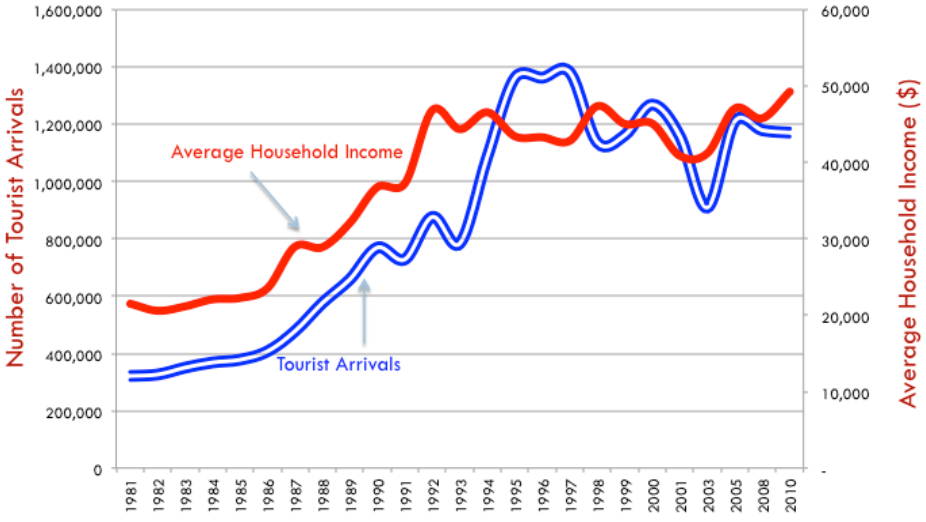
Empirical studies of Caribbean countries show a significant correlation between tourism industry and economic performance (Jayawardena and Ramajeesingh, 2003). Figure 2 presents the evolution of Guam's tourist arrivals and average household income from 1981 to 2010. The two curves reveal a general trend of increase in both tourist arrivals and average household income from 1981 to the early 1990s and a pattern of fluctuation in both thereafter.

Figure 1: Tourist Arrivals on Guam 1990-2011



Source: Guam Visitors Bureau

Figure 2. Tourist Arrivals vs. Average Household Income on Guam: 1981-2010

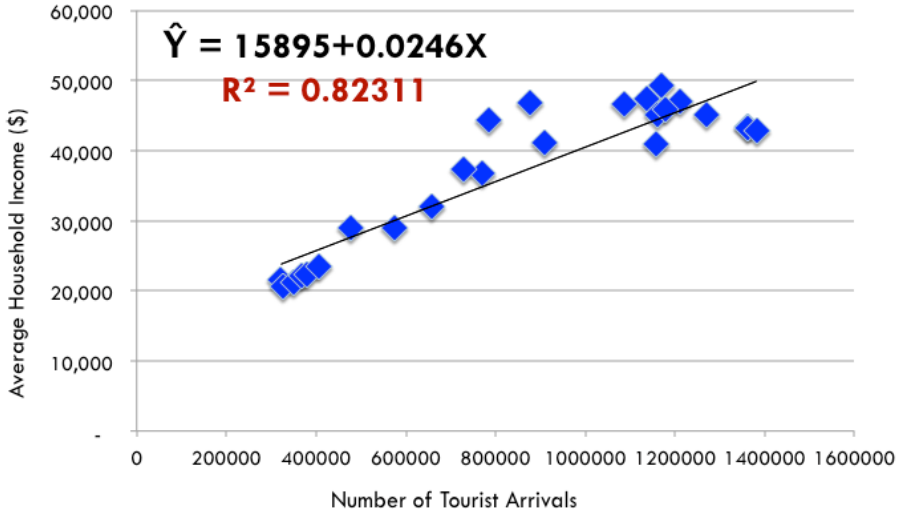


Source: Guam Visitors Bureau, Guam Department of Labor

In order to determine whether there is a linear relationship between average household income level and tourist arrivals on Guam, a simple regression model was constructed: $Y=\alpha+\beta X+\varepsilon$, where Y is average household income, X is number of tourist arrivals, and ε is the error term. The result of regression analysis shows a strong positive linear relationship between the two variables ($R^2=0.82,\ F=107,\ p<0.001$). Hence, number of tourist arrivals is a significant contributor to Guam's household income. In other words, over the years, the increase of tourist arrivals on Guam has resulted in increase in average household income. Figure 3 shows the scatter diagram and the estimated regress line.

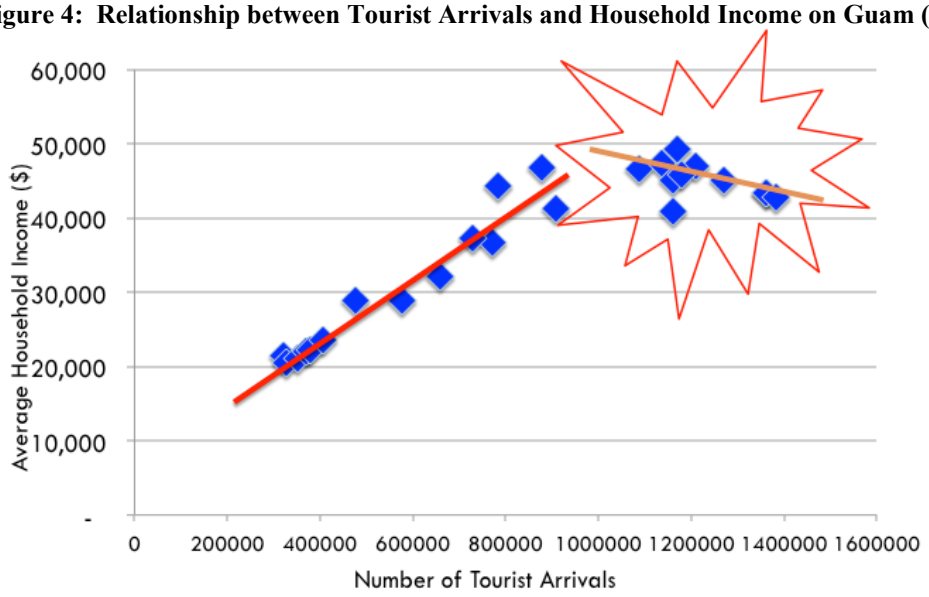
However, more insights are revealed if the scatter diagram is reorganized based on the number of tourist arrivals. Data points in the scatter diagram are grouped into two clusters: those with less than one million tourist arrivals and those with over one million. As shown in Figure 4, there is a strong positive linear relationship between tourist arrivals and average household income on Guam when Guam receives less than one million tourist arrivals ($R^2=0.94,\ F=184,\ p<0.001$), and the relationship presents a somewhat negative pattern after the number of tourist arrivals reaches one million ($R^2=0.25,\ F=3.05,\ p=0.11$). This finding is both interesting and surprising. Although it is not clear about the dynamics behind this negative cause-and-effect relationship, one thing seems certain: there is still much room for realizing the potential benefits of the tourism industry on Guam. Therefore, a call for new strategies is imperative.

Figure 3: Relationship between Tourist Arrivals and Household Income on Guam (a)



Sources: Guam Visitors Bureau, Guam Department of Labor

Figure 4: Relationship between Tourist Arrivals and Household Income on Guam (b)



Sources: Guam Visitors Bureau, Guam Department of Labor

Meyer (2006) argues that over the years an increasing number of tourist destinations have learned form their experience about the fallacy of paying too much attention on the number of tourist arrivals. Instead, the focus has been switched to increasing the benefits to the local economy. According to Meyer (2006), there are three strategies a small island state may use to increase benefits to the local economy: (1) "increasing spending per visitor"; (2) "increasing local participation in the industry"; and/or (3) "increasing backward linkages and thus reducing leakages." Here, the term "linkage" refers to the collaboration and usage of other business sectors on island so that to improve the local economy as a whole, and leakage is usually the percentage of the tourist spending that leaves or never reaches the island due to the involvement of intermediaries. The key point of these strategies is that the focus should be put on increasing the number of tourist arrivals, as well as on encouraging involvement of local business and on using local agricultural and other products for import substitution so that the benefits of tourist industry will remain on island. In Guam's case based on estimates by Global Insight, Inc., for every dollar spent by tourists in Guam, 34 cents go to the local tourism services or goods, 12 cents goes to local suppliers (linkage), and 54% goes to off-shore suppliers (leakage) (Pike, 2007). More than half of the tourist spending is lost due to import leakages, with unrealized benefits in the destination where goods and services are provided to the tourist. It appears that Meyer's three strategies could be implemented through one development model: the one-village-one-product (OVOP) model.

THE OVOP STRATEGY AS GUAM'S CHOICE TO IMPROVE THE TOURISM INDUSTRY

The OVOP movement was originated in Oita Prefecture of Japan as a regional development program in 1979, when Oita lagged behind the other parts of Japan in terms household income level. The movement was advanced by Oita's former governor Morihiko Hiramatsu. As governor, he advocated that communities should selectively produce goods with highly added value and one village should produce one product that was competitive and stable and use this particular product to gain sales revenue in the market. This movement was successfully implemented and eventually over 300 products were selected by villages in Oita. As a result, the average annual income in Oita doubled from $13,000 in 1979 to $26,000 in 2003 (Shoji, 2008). Today, the OVOP products are sold in Narita International Airport and Kansai International Airport, making revenues for villages involving in the OVOP movement.

The success of the OVOP movement has given other developing countries and regions hope to use it as a development strategy. Many developing countries, such as India, Egypt, Tanzania, Nepal, Myanmar, and Madagascar, have adopted the OVOP model for regional development. A few typical OVOP products from these countries include: different types of incense with aromas such as raspberry and strawberry (India), bookmarks featuring wall painting as reproductions of the oldest paper made in ancient times (Egypt), cellphone charms attached with animal and mask designs made of animal bones (Tanzania), Buddhist painting paper (Nepal), gripping stars made of natural wood as massage tool (Myanmar), and model bicycles made from waste cans (Madagascar). Each of these handicrafts is embedded in the village's unique social and cultural settings.

Although Guam has not adopted the OVOP strategy, there is strong evidence to show that Guam has great potential to benefit from such a development model and that of the OVOP model.

Because tourism serves as a major business sector and an engine for Guam's economic growth, it is necessary to combine the OVOP model with the development strategy of Guam's tourism industry. In fact, there has been a relatively strong connection between Guam's tourist industry and various village products in terms of agriculture, culture, history, and natural beauty. This is truly a unique feature favorable for the adoption of the OVOP strategy in Guam. Guam has 19 villages. In adoption of the OVOP model, by pursuing a specialization strategy, each village may identify one or two products (goods or services) that will eventually create a specific image to attract visitors and investments. For example, the OVOP model will help with import substitution strategy for agricultural goods by pursuing economies of scale. Therefore, the one village one product (OVOP) model seems to be a good strategy that Guam can adopt for its rural development and for the island's tourism industry. By doing so, Guam will be able to increase the linkage and reduce the leakage.

Guam's tourism products that normally attract visitors can be categorized into to the following:

- Natural Beauty: sea, sand, sun, cliff, and mountain
- Culture/History/People: Chamorro culture, Spanish heritage, and Chamorro people
- Sports: golf course, surfing, hiking, and camping
- Handicrafts: local arts, and hand-made decorations
- Agriculture: tropical fruits
- Lodging: 5-star hotels
- Food: restaurants
- Shopping: shopping centers and malls
- Wedding: wedding services
- Recreation: entertainment facilities

In order to take the OVOP model into consideration for Guam, it is necessary to determine whether at least one tourism product could be identified for each of the villages. Guam is divided into 19 municipalities called villages. Villages vary greatly in size measured by population, ranging from 782 in the village of Umatac to 44,943 in the village of Dededo. Table 1 lists the 19 villages on Guam and their population sizes.

Table 1: Guam's Villages and Populations

Villages	Population
Agana Heights	3,808
Agat	4,917
Asan	2,137
Barrigada	8,875
Chalan Pago-Ordot	6,822
Dededo	44,943
Hagåtña	1,051
Inarajan	2,273
Mangilao	15,191
Merizo	1,850
Mongmong-Toto-Maite	6,825
Piti	1,454
Santa Rita	6,084
Sinajana	2,592
Talofofo	3,050

Tamuning	19,685
Umatac	782
Yigo	20,539
Yona	6,480
GUAM TOTAL	159,358
Source: United States Census 2010	

Table 2 lists OVOP products candidates for almost every village on Guam. The table is compiled through investigation of each village's major points of interest. The purpose is to show that each village has some uniqueness in terms of its history, culture, agriculture, natural beauty, or facilities. Therefore, Guam has the foundation and potential to benefit from the OVOP model. Note that two villages, Chalan Pago-Ordot and Mongmong-Toto-Maite, are labeled question marks. No specific OVOP candidates were easily identified for these two villages, indicating further detailed investigation is required.

Table 2: OVOP Product Candidates in Guam's Villages

Villages	Candidate OVOP Products
Agana Heights	Fort Santa Agueda (history), Fonte River (sports), Coconut Theme Park (culture & agriculture)
Agat	Agat Beach (history), Agat Small Boat Marina (Recreation), Mango Festival (culture & agriculture)
Asan	Asan Beach (history: War in the Pacific National Historical Park)
Barrigada	Admiral Nimitz Golf Course (sports)
Dededo	Guam National Wildlife Refuge (natural beauty), Coco Palm Resort (natural beauty), Dededo Flea Market (culture & agriculture), Dededo Skate Park (sports), Micronesian Mall and Theaters (shopping), Stars Golf Course (sports), Tanguisson Beach Park (sports: hiking)
Hagåtña	Latte Stone Park (culture), Japanese World War II Caves (history), Plaza de España (history), Cathedral-Basilica (culture), Chamorro Village Night (culture and recreation)
Inarajan	Gef Pa'go Cultural Village (culture & history), Talofofo Bay Beach Park (sports: surfing)
Mangilao	Onward Mangilao Golf Club (sports), Academic Region (UOG and GCC)
Merizo	Cocos Island (natural beauty), Merizo Festival of the Sea (culture)
Piti	Fish Eye Marine Park (underwater observatory), Piti Bomb Holes Preserves (sports: swimming and diving), Mahogany Forest (agriculture)
Santa Rita	Namo Falls Tropical Garden (natural beauty and agriculture)
Sinajana	Ada's Funeral Home (culture)
Talofofo	Ipan Beach Resort (sports: camping and hiking), Jungle Riverboat Cruise (natural beauty), Talofofo Falls and Caves (natural beauty), Talofofo Golf Course (sports)
Tamuning	Guam Premier Outlets (shopping), Two Lover's Point, Tumon- Tourist District (lodging, shopping, entertainment), Ypao Beach Park (natural beauty and sports),
Umatac	Spanish Forts (history), Guam Veteran's Memorial (history)
Yigo	Mt. Santa Rosa (natural beauty: highest point in Northern Guam), Pacific Peace Memorial Park (history: Japanese Memorial for WWII)
Yona	Hamamoto Tropical Fruit World (agriculture), Windward Hills Country Club (sports), LeoPalace Resort (recreation)

CONCLUSIONS AND PUBLIC POLICY IMPLICATIONS

Its unique strategic location and the impending military build-up have offered Guam great opportunities for rapid economic growth in the near future. However, Guam still needs proper public policy and strategies to build up its indigenous capability to ensure sustainable development. One can identify two major business sectors on Guam--the military-related service sector and tourism. The former is plagued with a high level of uncertainty as the planned military buildup is still unclear in terms of scope and timing. The latter seems more promising as there is still much room for Guam's tourist industry to improve its performance with new strategies.

Over history, Guam's tourist industry has served as an engine for economic growth and has contributed to increasing income levels on Guam. This paper found a strong positive relationship between the number of tourist arrivals and the average household income. Research also revealed that when the number of tourist arrivals exceeds one million, the above positive relationship switches into negative. This may be an indicator that Guam is not fully benefiting from tourist spending as there is a high level of leakage. This finding is consistent with the findings of the study conducted by Global Insight, Inc. (Pike, 2007).

In order to promote the linkage and to reduce the leakage in tourism industry, this paper argues that the OVOP model seems to be a good strategy that Guam can adopt for its rural development and for the island's tourism industry. Investigation of each of Guam's 19 villages results in a list of candidate OVOP products for most of the villages, indicating Guam has potential to benefit form the OVOP strategy.

However, the OVOP strategy will not be successful without the support from public policy. It is recommended that the local governments on Guam need to work closely with agriculture and tourism industry officials to develop public policies that promote the OVOP movement. Specifically, in order to successfully implement the OVOP strategy for Guam's tourist industry, the Government of Guam needs to address the following (Sheldon, 2005; Shoji, 2008):

- **Consensus of the need to change**: Guam's tourism industry has been doing a good job in bring more visitors to Guam from different source markets. In recognition of the fact that this industry has a high level of leakage, there is a need to switch the focus to building up more linkages so that the local community will benefit more from tourist spending on Guam. A consensus of needing change and public support will ensure full involvement of the community in implementing new strategies towards the changing of focus.
- **Empowerment of the local community and culture:** The uniqueness of Guam's local culture and history is one of its major selling points to attract tourists. Providing tourists with the opportunities to enjoy and appreciate the local culture requires empowerment of the local community. Guam is currently

undergoing a cultural revival of sorts and must continue to make efforts to showcase its unique culture to both residents and tourists. This could done through local festivals, arts, language programs, story-telling, and memory of traditions (Sheldon, 2005).

- **Persistence with the OVOP strategy**: Once the OVOP strategy has been put in place, persistence becomes the key to success. Relevant government agencies should be assigned the task of measuring the effectiveness and efficiency of the OVOP strategy while constantly monitoring the implementation of the strategy.
- **Investigation to identify products (goods or services) of high value-added**: Identifying the OVOP product for a village means taking other products out of focus. Thus, careful investigation and thorough consideration are needed in making decision on the selection of the OVOP product. Products of high value-added should have higher priority.
- **Marketing and market diversification**: When marketing OVOP products (goods or services), aiming at generating the highest expenditure for the island, the focus should be on the quality visitors in terms of their expenditures and length of stay. A market diversification approach is also needed to attract visitors from various source markets and from different levels of income. In this way the profitability of OVOP products will be enhanced.
- **Human resource development**: Human resource or human capital is key to successfully implementing public policy and business strategy. Despite all the advances in technology to streamline business operations, tourism is a people-oriented industry that relies on properly trained people. Formal education and training programs are needed for people in the tourism industry as well as for general residents. The international tourism and hospitality program at the University of Guam, as well as other educational institutions, require continuous support to develop human resources. Various training programs should be planned and provided on a regular basis for Guam to maintain competitiveness as a destination.

With OVOP on Guam, residents will be implementing a strategy in an industry that to some extent (unlike the military buildup), Guam's residents have control and have the ability to sustain. In addition, it will be increasing employment opportunities for village residents, teaching skills to produce local products, and establishing pride in village and island products. Most importantly, it will be increasing opportunities for residents around the island to interact directly with tourists to sell their specialty products, thereby increasing linkage and minimizing leakage.

REFERENCES

Aguilo, E., J. Alegre, and M. Sard (2005). The persistence of the sun and sand tourism model. *Tourism Management, 26* (2), 219-232.

Armstrong, H. and R. Read (1998). Trade, competition and market structure in small states: The role of contestability, *Bank of Valletta Review 18*(1),1–18.

Benedict, B. (1967). Sociological aspects of smallness, in B. Benedict (ed.), *Problems of smaller territories* (pp. 45–55). London: Athlone Press.

Bertram, G. (2004). On the convergence of small island economies with their metropolitan patrons. *World Development, 32*(2), 343–364.

Bertram, G. and R.F. Watters (1985). The MIRAB economy in South Pacific microstates. *Pacific Viewpoint, 26*(3), 497–519.

Briguglio, L. (1995). Small island developing states and their economic vulnerabilities. *World Development, 23*(10), 1615–1632.

Brookfield, H.C. (1990). An approach to islands. In W. Beller, P. d'Ayala and P. Hein (eds.), *Sustainable development and environmental management of small islands* (pp. 23–33). Paris: UNESCO-Parthenon.

Chenery, H.B. (1960). Patterns of Industrial Growth. *The American Economic Review, 50*(4), 624-654

Chenery, H.B. and L. Taylor (1968). Development Patterns: Among Countries and Over Time. *The Review of Economics and Statistics, 50*(4), 391-416.

Demas, W.G. (1965) *The economics of development in small countries*. Montreal: McGill University Press.

Knox, A.D. (1967). Some economic problems of small countries. In B. Benedict (ed.), *Problems of smaller territories*. London: Athlone Press.

Jayawardena, C. and D. Ramajeesingh (2003). Performance of tourism analysis: a Caribbean perspective. *International Journal of Contemporary Hospitality Management, 15*, 176–179.

Kotler, P., M. A. Hamlin, I. Rein, and D. H. Haider. (2002). *Marketing Asian Places: Attracting Investment, Industry, and Tourism to Cities, States, and Nations*, Singapore: John Wiley & Sons (Asia) Pte Ltd

McElroy, J.L. and M. Mahoney (2000). The propensity for political dependence in island microstates. *International Journal of Island Affairs, 9*(1), 32–35

McElroy, J. L. (2006). Small island tourist economies across the life cycle. *Asia Pacific Viewpoint, 47*(1), 61–77.

Meyer, D. (2006). Caribbean tourism, local sourcing and enterprise development: Review of the literature. *PPT Working Paper* No. 18

Pike, C. (2007). *Guam Tourism Economic Impact: Preliminary Results*, Global Insight, Inc.

Rostow, W. W. (1960). *The Stages of Economic Growth: A Non-Communist Manifesto*, Cambridge University Press

Ruane, M.C. (2012) *First Hawaiian Bank 2011-12 Economic Forecast, Guam-CNMI Edition*, Retrieved May 5, 2012, from https://www.fhb.com/pdf/FHB_GuamCNMIEconForecast2011.pdf

Sheldon, P. J. (2005). *The Challenges to sustainability in island tourism*. School of Travel Industry Management, University of Hawaii, Occasional Paper, 2005-1

Shoji, K. (2008). "The 'One Village, One Product' Movement Spreading throughout the World", Speech on OVOP APEC 2008, Retrieved April 30, 2012, from http://www.apecovop.org/ebs01-1.asp?page=&forid=4&todir=15&nouse=2247

A THEORETICAL ANALYSIS OF ISO9000 SUPPLIERS

Chiaho Chang, Montclair State University

ABSTRACT

This paper looks at the economic behavior of suppliers under different ISO9000 standards, especially that of ISO9001 and ISO9002. Considering the information asymmetry, given the same quality provided, the cost-reducing efforts of the ISO9001 SUPPLIER and ISO9002 SUPPLIER under a fixed, cost-plus contract are investigated. The result shows that the cost-reducing effort of the ISO9002 SUPPLIER is in line with the main manufacturer's expectations while the ISO9001 SUPPLIER is able to keep some of the information rent and exerts less effort. The bargaining power of the ISO9001 SUPPLIER is also stronger relative to that of the ISO9002 SUPPLIER. It is consistent with the degree of information asymmetry residing in the contracting parties. This paper also derives an optimal contract and procurement policy based on a simple institutional setting. The optimal cost-plus contract is obtained and its components are deciphered. The drivers behind the supplier's cost-reducing effort are also studied.

Outsourcing is the strategic use of outside resources to perform activities traditionally handled by internal staff and resources. It has been practiced for years, but the recent surge in excitement and growth is likely to result from changes in the competitive marketplace, which force the companies to take a hard look at their core competencies and form a closer alliance with their suppliers to help reduce costs and improve services. Supply chain management, as part of the enterprise resource planning (ERP) paradigm, becomes critical for survival.

The success of the Japanese auto makers has generated significant interest for researchers to look into many aspects of the Japanese style of management. On operational side, important concepts in Japanese production system such as kanban system, zero inventory, and just-in-time have been formalized and heavily studied (e.g., Monden, 1983; Hall & Hall, 1984). This development has also ignited research in analyzing the impact of setup cost reduction on production planning (e.g., Porteus, 1985; Zangwill, 1987). When US manufacturers encountered difficulties in implementing kanban or just-in-time system, researchers began to study the behavioral side of the Japanese system, especially corporate culture such as business groups and keiretsu, lifetime employment, and team work (e.g., Hutchins, 1986; Abegglen & Stalk, 1985; Imai, 1986). However, when it comes to one of the building blocks of the Japanese auto industry - the suppliers, the incentive issues as applied to the relationship between the assembler and its suppliers receive little attention.

One major reason is that the data itself is difficult to come by. Researchers are not able to get a hold of the details of contract negotiations and the finalized version of the contract itself. Even if they are familiar with the operations, the research methodology used is based on case-by-case, descriptive field study, which is difficult to generalize. Second, even written contracts sometimes are vague. Informal, implicit agreements constitute a large portion of these black-box elements. Third, most researchers in this area come from disciplines such as operations research, industrial engineering, management science and organizational behavior which usually regard such supplier relationship as being smooth, thereby assuming away the incentive problems. Economists are indeed interested in optimal incentive schemes. However, their derivations are usually done without taking into account what practice dictates, a criticism rightfully advanced by Arrow (1985, 48).

Asanuma (1985a, 1989) has conducted extensive field studies in the Japanese auto industry. Three sources of components were identified. They are (1) design approved (DA), where the supplier provides both the manufacturing capability and technical know-how for the design approved by the assembler; (2) design supplied (DS), where the assembler provides the technical drawings and the supplier provides only the manufacturing capability; and (3) off the shelf (OS) for standard components. In the realm of supply chain management, the assembler is most interested in the first two sources, which will be called ISO9001 SUPPLIER and ISO9002 SUPPLIER in this paper for generality.

This paper tries to elicit and contrast the cost-reducing efforts of the ISO9002 SUPPLIER and ISO9001 SUPPLIER given a fixed, cost-plus contract. Reducing and controlling operating costs is listed as the top reason companies outsource. As the practice of target costing spreads rapidly, the assembler has every intention of "... transmit[ting] the competitive reality faced by the firm to its suppliers." (Cooper & Slagmulder, 1997, 14) How the suppliers react to the assembler's demand and whether there is any difference among suppliers become interesting issues. In addition, this research derives an optimal contract and procurement policy based on a simple institutional setting in order to address the incentive issues involved in observed Japanese practice of adopting linear contracts (Asanuma, 1985a) in such relationships.

The remainder of the paper is organized as follows. The next section provides literature review on related issues. This is followed by an analysis of the supplier behavior given a fixed, cost-plus contract, as well as the derivation of the optimal procurement policy. The last section concludes the paper.

LITERATURE REVIEW

Although target costing is usually classified as one form of market-based pricing, its value as a cost-reducing tool cannot be overemphasized. Once set, "[t]*he target cost of a product can never be exceeded*." The firm then uses techniques such as value engineering and quality

function deployment to modify design, material specification and production process to reduce costs while preserving the value as perceived by the customers. For outsourced components, the assembler transmits the market pressure to the supplier in the form of target price paid, which in turn becomes the supplier's target cost to meet.

Loeb & Surysekar (1998) study whether and how payment ceilings should be set in cost-plus contracting. Their findings support the use of an overall payment ceiling to elicit the supplier's private (cost) information and to mitigate the moral hazard problem associated with cost-plus contracting. However, when target costing paradigm is adopted, as is done in this paper, both "whether" and "how" problems with respect to payment ceilings become moot at best. The ceiling is already determined by the market conditions.

Laffont & Tirole (1986) consider a static (one-period) control problem where a regulated firm with private information about its own efficiency parameter decides how much level of effort to put into production process. The regulator (e.g., government agency) has prior belief on the firm's "type" and observes the actual cost of production. They are able to derive an optimal scheme which is linear in ex post cost. See also Holmstrom & Milgrom (1987) for similar results.

In Laffont & Tirole (1988), they preserve most of the basic structure of their 1986 paper, including the efficiency parameter, but extend it to a dynamic (two-period) framework. This paper formalizes the concept of "ratchet effect" by allowing the regulator not to "commit himself not to use in the second period the information conveyed by the firm's first-period performance."

The problem with the use of the efficiency parameter to identify the type of the firm is that when there are more than two periods, as the models in this paper adopt, it becomes difficult to update the regulator's belief reasonably well unless an appropriate equilibrium concept is invoked such as sequential equilibrium (Kreps & Wilson, 1982), an exercise not tried here. So instead the efficiency parameter is replaced by a random variable that represents the unpredictable production environment (e.g., how likely the machinery will break down or the yield rate of the output) against which the supplier exerts effort to tame the cost of production. The realization of this random variable is observable only by the supplier before she makes effort decision but the assembler has some preliminary information about it (i.e., knows its probability distribution).

Another problem with Laffont & Tirole (1988) is that, although a two-period model provides sharper focus and tractability, it simply cannot capture the long-term relation between the assembler and his suppliers. The multi-period models presented in this paper thus subsume the two-period one and eventually are extended to infinite horizon.

The cost structure used in this paper is similar to that of McAfee & McMillan (1986), with two differences. First, it is indexed by time in a multi-period setting; second, the target cost at period t replaces the intrinsic cost that is observable only by the supplier. They also compare an incentive contract with cost-plus and fixed-price contracts in a bidding situation and conclude that the incentive contract performs better. Since their model is essentially one-period, target

cost plays no role except in the trivial case where average cost is calculated from previous periods. Kawasaki & McMillan (1987) then use their results to empirically examine the parameters of the incentive contract in the context of subcontracting in Japanese manufacturing industries. It is tempting to use the incentive contract because of strong empirical implications. But as Asanuma (1985a) points out, the contracts between the assembler and his suppliers are basically cost-plus. So the efficiency issue of the incentive contract will be put aside for future studies

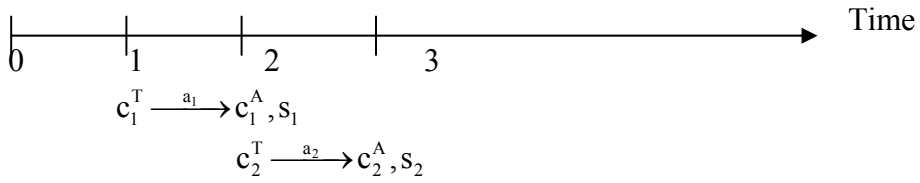
As to the ratchet effect, Weitzman (1980) provides an early treatment on this topic. He models a no-commitment situation by explicitly formulating target output as a function of the agent's previous performance, as is done in this paper. But he treats the parameters of the target as random variables. Instead, this paper leaves these parameters fixed, as is determined in the negotiation process before mass production begins, so the bargaining power of the parties to the contract can be examined.

Recent development in the literature casts the issue of specific investment (or reliance investment in contract law jargon), such as the cost-reducing effort in this paper, in the realm of incomplete contracts and renegotiation (e.g., Chung, 1991; Hart & Moore, 1988; Reichelstein, 1992). Gietzmann & Larsen (1998) study how cooperation between the assembler and the supplier can be forthcoming via a careful design of the governance procedures in an incomplete contract setting. Since the parameters of the contract considered in the model are assumed fixed ex ante, such complexity is avoided.

SUPPLIER BEHAVIOR GIVEN FIXED CONTRACT

In this section, the supplier's cost-reducing behavior given a fixed, cost-plus contract will be extracted. The model considered has two pairs of players: the assembler will be matched with the ISO9002 SUPPLIER and the ISO9001 SUPPLIER, respectively. They are all assumed to be risk neutral in order to focus on incentive issues. The assembler signs contracts with the two types of suppliers following the contractual scheme in Figure 1.

Figure 1



During time 0, a negotiation is initiated by the assembler to determine the parameters of the target costs (c_t^T) and the incentive payments (s_t) to be used for all future periods (t = 1, 2, ..., T). At the beginning of each period t, the assembler and the supplier will compare previous period's actual cost (c_t^A) with the target cost, settle the payment, and determine current period's target cost. During each period t, the supplier will then contribute her effort (a_t) to reduce cost as postulated by the target cost.

The target cost at time t is indirectly determined by

$$c_{t-1}^T - c_t^T = \delta + \lambda(c_{t-1}^T - c_{t-1}^A), \tag{1}$$

where $\delta > 0$ is the fixed cost-reducing goal over the length of the contract and represents how much the cost should be reduced in period t if last period's target cost were exactly met. $\lambda \in [0,1]$ is the adjustment parameter. Both δ and λ are determined ex ante at time 0. This formula has the supplier's continuous improvement effort built into the contract. (1) can be rewritten (1) as

$$c_t^T = \lambda c_{t-1}^A + (1-\lambda)c_{t-1}^T - \delta. \tag{2}$$

It is easy to see that period t's target cost can be expressed as a weighted average of period t-1's actual and target costs, less a fixed cost-reducing goal. Given c_0^A and c_0^T as initial values, (2) can be further rewritten as

$$c_j^T = c_0^T(1-\lambda)^j + \sum_{t=0}^{j-1}(\lambda c_t^A - \delta)(1-\lambda)^{j-t-1}.$$

The actual cost at time t, c_t^A, can be denoted as

$$c_t^A = c_t^T + w_t - a_t \tag{3}$$

and is observable ex post to the contracting parties. w_t is an i.i.d. random variable at time t, representing unpredictable cost fluctuations whose realization is observed only by the ISO9001 SUPPLIER during the manufacturing process, but the assembler has a prior belief of w_t, $f(w_t)$, defined over the interval $[\overline{w}, \underline{w}]$, a fixed support. As to that of the ISO9002 SUPPLIER, it is assumed that there is no information asymmetry and the assembler is able to observe its realization with certainty. a_t represents the extent to which actual costs are reduced as a result of the supplier's effort. It can also be interpreted as the relation-specific investment made by the supplier.

With a cost-plus contract, it can be assumed that the gross and net payments from the assembler to the supplier are, respectively,

$$g_t = s_t + c_t^A$$

and

$$s_t = k + \alpha(c_t^T - c_t^A),$$

where $k > 0$ is the gross profit margin and $\alpha \in [0,1]$ the reward parameter. k and α are determined ex ante. This format is in spirit similar to Laffont & Tirole's (1986) result: a contract linear in ex post cost.

The supplier's utility function, in monetary terms, is

$$u_t = s_t - H_t(a_t),$$

where $H_t(a_t)$ is the supplier's cost of effort. It is assumed to be increasing and convex (i.e., $H_t' > 0$ and $H_t'' > 0$). A common discount factor is assumed for all parties: γ.

The assembler's problem with respect to the ISO9002 SUPPLIER supplier can be described as follows:

$$\min_{\{a_t, s_t\}_{t=1}^T} \sum_{t=1}^{T} \gamma^{t-1} g_t \qquad \text{(A-ISO9002 SUPPLIER)}$$
$$\text{s.t. } u_t \geq 0, \forall t.$$

The assembler wants to minimize his total discounted payment over T periods subject to the ISO9002 SUPPLIER receiving at least a reservation level of utility (normalized to zero). Since this problem of perfect information is a stationary one, the assembler is in effect solving, for each period,

$$\min_{\{a,s\}} g$$
$$\text{s.t. } u \geq 0.$$

Proposition 1

Without information asymmetry, the optimal contract between the assembler and the ISO9002 SUPPLIER can be characterized by

$$u = 0$$

and

$$H'(a^{OEM}) = \alpha = 1 .$$

Proof (Omitted)

In this problem, the ISO9002 SUPPLIER will receive only her reservation utility and exert a level of effort that is Pareto efficient because of symmetry of information.

Next, consider the assembler's problem when he faces an ISO9001 SUPPLIER:

$$\min_{\{a_t, s_t\}_{t=1}^{T}} \sum_{t=1}^{T} \gamma^{t-1} \int_{\underline{w}}^{\bar{w}} g_t f(w_t) dw_t \qquad \text{(A-ISO9001 SUPPLIER)}$$

$$\text{s.t. } u_t \geq 0, \forall t \text{ and } \{a_t\}_{t=1}^{T} \text{ maximizes } V = \sum_{t=1}^{T} \gamma^{t-1} u_t .$$

The revelation principle does not apply here in the absence of commitment. Moreover, the assembler is not concerned about the ISO9001 SUPPLIER supplier's report on w_t any more than her cost-reducing effort. So the assembler will minimize his total expected discounted payments subject to the ISO9001 SUPPLIER's individual rationality and incentive compatibility constraints.

Given the passive target-setting role of the assembler, a set of optimal decision rules $\{a_t^{ODM}\}$ for the ISO9001 SUPPLIER can be found by solving her decision problem alone. The optimal solution can be described in the following proposition.

Proposition 2

The optimal contract between the assembler and the ISO9001 SUPPLIER can be characterized by

$$H_t'(a_t^{ODM}) = \frac{\alpha}{1 + \frac{\lambda\gamma}{1-\gamma}} \leq \alpha .$$

Proof

The ISO9001 SUPPLIER's objective function can be expressed as

$$V = \sum_{j=1}^{T} \gamma^{j-1} u_j \quad . \tag{4}$$

To have a closed-form solution, let $T \to \infty$ [12] and use the fact that

$$\sum_{j=t+1}^{\infty} \gamma^{j-1}(1-\lambda)^{j-t-1} = \frac{\gamma^t}{1-\gamma+\lambda\gamma} \quad .$$

Then (4) can be rewritten as

$$V_\infty = \sum_{t=1}^{\infty} \gamma^{t-1}[k + \alpha(c_0^T(1-\lambda)^t - c_t^A) - H_t(a_t)] + \sum_{t=0}^{\infty} \alpha(\lambda c_t^A - \delta)\frac{\gamma^t}{1-\gamma+\lambda\gamma} \quad , \tag{5}$$

where V_∞ indicates that an infinite horizon problem is being solved.

Using (3), (5) can be reduced to

$$V_\infty = \sum_{t=1}^{\infty} \gamma^{t-1}\left[\frac{\alpha}{1+\frac{\lambda\gamma}{1-\gamma}} a_t - H_t(a_t)\right] + Z \quad , \tag{6}$$

where

$$Z = \sum_{t=1}^{\infty} \gamma^{t-1}\left[k + \alpha c_0^T(1-\lambda)^t - \frac{\alpha}{1+\frac{\lambda\gamma}{1-\gamma}}(c_t^T + w_t)\right] + \alpha\lambda c_0^A \frac{1}{1-\gamma+\lambda\gamma} - \sum_{t=0}^{\infty} \alpha\delta \frac{\gamma^t}{1-\gamma+\lambda\gamma} \quad .$$

Note that Z is a constant independent of $\{a_t\}$. The variable part of (6) is additively separable across periods in functions of a_t. Therefore, (6) will be maximized if and only if in each period t, a_t is selected to maximize

$$\frac{\alpha}{1+\frac{\lambda\gamma}{1-\gamma}} a_t - H_t(a_t) \quad ,$$

or

$$H_t'(a_t^{ODM}) = \frac{\alpha}{1+\frac{\lambda\gamma}{1-\gamma}} \le \alpha \quad .$$

Note that the optimal value, a_t^{ODM}, does not depend on c_t^T. Given that the second-order condition ($H_t'' > 0$) is satisfied by assumption, the optimal value must be an interior solution. *Q.E.D.*

Overall, the solutions seem myopic at best. Both types of the suppliers will only look at the parameters negotiated at time 0 to determine their behaviors. In the case of (A-ISO9001 SUPPLIER), where the assembler has imperfect information about w_t, the supplier will be able to exert less effort and enjoy more information rent than in the case of (A-ISO9002 SUPPLIER), where the assembler has complete control. To induce more effort, the assembler has to reward more (i.e., increase α) and/or punish less (i.e., decrease λ) for the ISO9001 SUPPLIER's investment in cost-reducing effort.

The solution to (A-ISO9002 SUPPLIER) says nothing about λ with respect to the ISO9001 SUPPLIER. Presumably, it should be higher than that for the ISO9001 SUPPLIER to bring the ISO9002 SUPPLIER in line with the assembler's policy. It can be called a "carrot-and-stick" approach toward the ISO9002 SUPPLIER.

On the other hand, since these parameters are determined ex ante during the negotiation process before mass production begins, this scheme calls for more bargaining power for the ISO9001 SUPPLIER as opposed to the ISO9002 SUPPLIER, relative to that of the assembler. This may be called a "honey-and-sugar" policy for the ISO9001 SUPPLIER.

It seems paradoxical at first to compare the results of Propositions 1 and 2 because the ISO9002 SUPPLIER receives $H_t'(a_t^{OEM}) = \alpha = 1$ while the ISO9001 SUPPLIER receives $H_t'(a_t^{ODM}) \leq \alpha$, implying that the ISO9002 SUPPLIER may be given a better bargaining position in terms of α. In fact, the larger share of (relation-specific) investment gain paid to the ISO9002 SUPPLIER can be interpreted as merit from the assembler and his intention to cultivate the ISO9002 SUPPLIER, who is more vulnerable, rather than an expression of larger bargaining power on the part of the ISO9002 SUPPLIER.

OPTIMAL PROCUREMENT POLICY

The detailed contractual relationship between the assembler and the supplier will be studied in this section.

In his procurement policy, the assembler specifies x units of a particular component for trial production and z units of projected demand for mass production by the designated supplier, if this stage is ever reached. The supplier realizes the unit cost of c during trial production and reports $\hat{c}$ instead.

Unit compensation is a function of the reported cost, $s = s(\hat{c})$ and is agreed upon beforehand. It is assumed that s is increasing in $\hat{c}$. The assembler has a procurement target cost, c^T, that he is willing to pay for each unit of the component acquired. This target cost is

determined by considering competitive price in the market and his own desired target profit and is given exogenously in the model.

The final target unit price for mass production, p^T, is determined by

$$p^T = \min \{s(\hat{c}), c^T\}.$$

The rationale is simple. If the supplier reports $\hat{c}$ such that the required compensation $s(\hat{c})$ is smaller than c^T, the assembler will not insist on paying c^T. That is, $p^T = s(\hat{c})$ for $s(\hat{c}) \le c^T$. On the other hand, if $s(\hat{c})$ is larger, there will be no mass production allowed unless the supplier accepts a lower compensation, c^T. That is, $p^T = c^T$ for $s(\hat{c}) > c^T$.

In order to produce the x units (and the z units, if called for later) of the component, the supplier has to invest a fixed cost F in capital assets, which will also enable her to find out the actual cost of production. A portion of the fixed cost, $(1-\beta)F$, can be recovered if no mass production follows. In other words, βF can be regarded as sunk once trial production begins.

In this section, a long-term relationship between the assembler and the supplier exists when the mass production is conducted following the trial production. Short-term relationship, on the other hand, indicates a situation where the supplier quits after just the trial production.

THE SUPPLIER'S PROBLEM

The model is developed backward from the mass production stage on. At mass production stage, it is assumed that a price target, p^T, has been agreed upon. Then the supplier has to choose an optimal level of cost-reducing effort, a^*, to maximize

$$V^P = \int_0^N \{p^T - [c - B(a)]\} z e^{-\gamma t} dt - \phi a,$$

where

V^P = the net present value of the supplier's profit during mass production, evaluated at time 0, when mass production begins (see Figure 2),
ϕ = the acquisition cost per unit of cost-reducing effort,
γ = the discount rate, or the supplier's cost of borrowing funds elsewhere,
N = the length of the mass production period,
a = the number of units of cost-reducing effort, and
B(a) = cost saving per unit time upon adoption of a units of cost-reducing effort, $B'(a) > 0$.

Recall that c and z are, respectively, the supplier's realized unit cost of production and mass production volume. The following proposition summarizes the supplier's optimal responses at mass production stage.

Proposition 3

(1) Increases in cost-reducing effort reduce production costs at a diminishing rate.
(2) Cost-reducing effort will not be undertaken unless the supplier is allowed to at least recover her costs.
(3) Increases in the length of the mass production period encourage more cost-reducing effort.
(4) Higher cost of borrowing funds will lower the optimal level of the cost-reducing effort.
(5) Increases in the cost of cost-reducing effort decrease the optimal level of the cost-reducing effort.

Proof

V^P can be rewritten as

$$V^P = \frac{1}{\gamma}(1-e^{-\gamma N})\{p^T - [c - B(a)]\}z - \phi a .$$

Necessary and sufficient conditions for the supplier's problem are, respectively,

$$V^{P'} = \frac{dV^P}{da} = \frac{1}{\gamma}(1-e^{-\gamma N})B'z - \phi = 0 \tag{7}$$

and

$$V^{P''} = \frac{d^2V^P}{da^2} = \frac{1}{\gamma}(1-e^{-\gamma N})B''z < 0 \tag{8}$$

From (8), we know that $B'' < 0$. In other words, cost savings are increasing at a decreasing rate as the level of cost-reducing effort increases.

If $N = 0$, then $V^P = -\phi a$. In this case, the optimal solution will have the supplier exert no cost-reducing effort; i.e., $a = 0$. Comparative statistics results are derived from (7) using implicit function rule.

$$\frac{da}{dN} = \frac{-\gamma e^{-\gamma N}B'}{(1-e^{-\gamma N})B''} > 0 .$$

$$\frac{da}{d\gamma} = \frac{[1-(1+\gamma N)e^{-\gamma N}]B'}{\gamma(1-e^{-\gamma N})B''} < 0 .$$

$$\frac{da}{d\phi} = \frac{\gamma}{(1-e^{-\gamma N})B^{*}z} < 0 \quad .Q.E.D.$$

This proposition shows the possibility to implement a lagged price adjustment scheme in which the assembler sets a price which will last for a certain period of time (in this model, N periods) and allows the supplier to exert cost-reducing effort and enjoy cost savings therein. However, it is only partially implemented because in this model there is no review of target cost after mass production begins and therefore no new (lower) target cost being set. The results are still valid and can provide policy guidance for the assembler. For example, to encourage cost-reducing effort, the assembler can extend the mass production period, arrange low-cost funds for the supplier, or even make the supplier's effort less costly by providing technical assistance.

Next, assume that the supplier wants to maintain long-term relationship with the assembler. To formalize this idea, let the supplier choose $\hat{c}$ so that $s(\hat{c})$ satisfies

$$V^{S}(a^{*}) + sx - F - cx \geq sx - \beta F - cx, \tag{9}$$

where

$$V^{S}(a^{*}) = \max_{a} V^{S}\left(= \int_{0}^{N} \{s(\hat{c}) - [c - B(a)]\} z e^{-\gamma t} dt - \phi a\right).$$

The left-hand side of the inequality (9) represents what the supplier will receive, evaluated at time 0, if she participates in mass production when $s(\hat{c})$ is paid; the right-hand side, her exit compensation from trial production. (9) can be simplified to get

$$V^{S}(a^{*}) \geq (1-\beta)F. \tag{10}$$

In other words, it is assumed that the supplier will not exaggerate reported cost of production "too much" in order to earn higher short-term profit from trial production and quit afterwards. From this assumption follows the next proposition.

Proposition 4

(1) There exists a critical value, $\underline{s}$, such that the supplier will participate in mass production only if $s(\hat{c}) \geq \underline{s}$.

(2) In the model, a full-cost-plus compensation scheme is necessary to sustain the long-term relationship between the assembler and the supplier.

Proof

From (10), by solving explicitly for $s(\hat{c})$, a critical value, $\underline{s}$, can be found such that

$$s(\hat{c}) \geq \underline{s} = [c - B(a^*)] + \frac{\gamma}{1 - e^{-\gamma N}} \frac{(1-\beta)F + \phi a^*}{z} . \quad (11)$$

It is consistent with a full-cost-plus contract as is normally observed in practice. To see why, express $s(\hat{c})$ as

$$s(\hat{c}) = \underline{s} + k_1$$

or

$$s(\hat{c}) = \underline{s}(1 + k_2), \quad (12)$$

where $k_1 \geq 0$ is the profit margin and $k_2 \geq 0$ the profit margin ratio. It is obvious that both equations in (12) satisfy (11) and are indeed full-cost-plus contracts desired by the supplier. *Q.E.D.*

The right-hand side of (11) indicates that, from the supplier's perspective, the assembler should pay, for each unit produced, at least the cost of production less the cost savings achieved ($c - B(a^*)$) plus the compensation for part of the fixed cost ($(1-\beta)F$, which would have been recovered from quitting after trial production) and cost-reducing effort (ϕa^*), both unitized by the mass production volume and multiplied by a time factor.

Notice that another part of the fixed cost, βF, is missing from the formula. It is tempting to interpret this as having been sunk already, with or without mass production. However, another interpretation for its absence in the critical value formula may be more plausible in this setting and has a counterpart in real-world situation. That is, it may be composed of capital outlays for equipment such as dies and tools which have alternative uses for the supplier in other projects. Since it is not specifically related to the assembler's project, the supplier does not expect to get reimbursed for such expenditures.

Since $s(\hat{c})$ is increasing in $\hat{c}$ by assumption, it is invertible. From (11), we can also find a critical value for $\hat{c}$, $\underline{c}$, such that

$$\hat{c} \geq \underline{c} = s^{-1}(\underline{s}) . \quad (13)$$

Proposition 5

To maintain long-term relationship with the assembler and remain viable, the supplier will report cost of production satisfying (13).

Corollary

Whether the supplier reports the true cost or not is irrelevant in this model. Trying to induce the supplier to report the true cost of production may not be efficient.

Proof

The proof is done through a counterexample. Consider the case where the supplier underreports cost of production (i.e., $\underline{c} \le \hat{c} \le c$) in order to launch mass production and recoup losses later through cost-reducing effort. If she is forced to tell the truth, mass production may never get started because of the assembler's target cost constraint. Both parties suffer. *Q.E.D.*

THE MAIN MANUFACTURER'S PROBLEM

Designate the optimal value of V^P (the net present value of the supplier's profit during mass production, evaluated at time 0) by $V^P(a^*)$. Then it becomes obvious that, from the assembler's perspective, the supplier will participate in mass production only if she cannot do worse participating than simply pulling out after trial production. That is, the following condition must be satisfied:

$$V^P(a^*) \ge (1-\beta)F .$$

Equivalently, a sufficient condition for the supplier's departure is

$$V^P(a^*) < (1-\beta)F . \qquad (14)$$

(11), (12) and (13) together shows that the sufficient condition for the supplier's departure can be extended to

$$s(\hat{c}) \ge \underline{s} > p^T = c^T ,$$

or

$$\hat{c} \ge s^{-1}(\underline{s}) > s^{-1}(p^T) = s^{-1}(c^T) .$$

One the one hand, the supplier asks for at least $\underline{s}$ for each unit produced, taking into consideration cost savings potential from optimal cost-reducing effort exerted; on the other, constrained by target procurement cost, the assembler wants to pay less than that, effectively asking the supplier to exert more effort (than she is willing to). In this case, no agreement can be reached and the supplier's departure becomes inevitable.

If the supplier decides to quit, the assembler's project may be in jeopardy. In the model, there is no obvious way out unless it is extended. One possibility is to introduce a second qualified supplier, the timing of which can be either at the beginning of trial production stage or after breakdown of negotiation. The first case allows for competition and presumably will lower the target price, thereby bringing it under the cap. The second case takes advantage of the first supplier's reported cost ($\hat{c}$), which is publicly available, and allows the assembler to invite only those qualified suppliers who are willing to produce the z units at a cost less than $\hat{c}$.

A second possibility is to negotiate a long-term contract with the supplier whose duration will cover several mass production stages. In such a contract, the supplier will be asked to stay throughout the whole contract period and accept pre-defined cost-reducing targets over time in exchange for initial higher compensation. This way, the assembler will break even or do better, depending on his target cost goal.

CONCLUSION

When confronted with increasing pressure to lower costs, a (utility maximizing) supplier will react to the assembler's contract offer with corresponding level of cost-reducing effort exerted, given that the contract is accepted. However, different types of suppliers are expected to react differently. Considering the Japanese automotive industry in particular and the manufacturing businesses in general as the backdrop, this paper compares the behavior of the ISO9002 SUPPLIER and the ISO9001 SUPPLIER in the presence of a fixed, cost-plus contract. The different degrees of information asymmetry between the assembler and the two suppliers lead the former to have complete control over the ISO9002 SUPPLIER supplier's cost-reducing effort while leaving the ISO9001 SUPPLIER room for information rent. The issue of bargaining powers between contracting parties is also explored.

The contractual scheme is then relaxed to derive an optimal procurement policy for the assembler. It turns out to be a linear one, the transfer payment to the supplier consisting of net production cost (i.e., production cost net of savings from cost-reducing effort) plus compensation for the costs of cost-reducing effort and part of the fixed assets purchased for the project. The result also demonstrates the potential to implement a lagged price adjustment mechanism in which the supplier enjoys additional cost savings once the target cost has been met during the current contract period. In other words, the extra savings from the supplier's cost-reducing effort will not be exploited by the assembler until the next round of contract negotiation begins, in which a new (and lower) target cost will be set. The assembler is encouraged to foster a closer tie with the supplier through longer-term relationship building, providing technical and technological assistance, and even arranging lower-cost loans for the supplier in exchange for the latter's willingness to reduce costs further. The assembler will be better able to share market pressure with his network of suppliers and concentrate on improving products and services.

Two limitations to the modeling approach here can be relaxed or amended in future research. The assembler's target procurement cost plays a crucial role in determining the fate of the mass production stage and the project as a whole, but it is given exogenously. It would be better if this target cost can be determined as a decision variable in the model. Also, the model entails essentially one (big) period, leaving price adjustment incomplete and the assembler's role passive. Extending the model to one more period will infuse richer results.

REFERENCES

Abegglen, J.C. & Stalk, G. (1985). *Kaisha, The Japanese Corporation*. New York, NY: Basic Books.

Anton, J.J. & Yao, D.A. (1987). "Second Sourcing and the Experience Curve: Price Competition in Defense Procurement." *RAND Journal of Economics*, 18(1), 57-76.

Arrow, K. J. (1985). "The Economics of Agency" in *Principals and Agents: The Structure of Business* ed. by Pratt, J.W. & Zeckhauser, R.J. Boston, MA: Harvard Business School Press, 37-51.

Asanuma, B. (1985a). "The Organization of Parts Purchase in the Japanese Automotive Industry." *Japanese Economic Studies* (Summer), 32-53.

Asanuma, B. (1985b). "The Contractual Framework for Parts Supply in the Japanese Automotive Industry." *Japanese Economic Studies* (Summer), 54-78.

Asanuma, B. (1989). "Manufacturer-Supplier Relationships in Japan and the Concept of Relation-Specific Skill." *Journal of the Japanese and International Economy*, 3, 1-30.

Blocher, E.J., Chen, K.H. & Lin, T.W. (1999). *Cost Management: A Strategic Emphasis.* New York, NY: McGraw Hill.

Christopher, M. (1998). *Logistics and Supply Chain Management*. London, UK: Prentice Hall.

Chung, T. (1991). "Incomplete Contracts, Specific Investment, and Risk Sharing." *Review of Economic Studies*, 58, 1031-1042.

Cooper, R. (1993). *Citizen Watch Company, Ltd.* Harvard Business School Case 9-194-033. Cambridge, MA: Harvard Business School.

Cooper, R. & Slagmulder, R. (1997). *Target Costing and Value Engineering.* Portland, OR: Productivity Press.

Davis, C.E., Davis, E.B. & Moore, L.E. (1998). "Outsourcing the Procurement-through-Payables." *Management Accounting* (July), 38-44.

Demski, J.S., Sappington, D.E.M. & Spiller, P.T. (1987). "Managing Supplier Switching." *RAND Journal of Economics*, 18(1), 77-97.

Farrell, J. & Gallini, N.T. (1988). "Second-Sourcing as a Commitment: Monopoly Incentives to Attract Competition." *The Quarterly Journal of Economics* (November), 673-694.

Gietzmann, M.B. & Larsen, J.G. (1998). "Motivating Subcontractors to Perform Development and Design Tasks." *Management Accounting Research*, 9, 285-309.

Hall, A. & Hall, R. (1984). *Zero Inventories*. Homewood, IL: Irwin.

Hart, O. & Moore, J. (1988). "Incomplete Contracts and Renegotiation." *Econometrica*, 56, 755-785.

Holmstrom, G. & Milgrom, P. (1987). "Aggregation and Linearity in the Provision of Intertemporal Incentives." *Econometrica*, 303-328.

Horngren, C.T., Foster, G. & Datar, S.M. (2000). *Cost Accounting: A Managerial Emphasis*. 10th Ed. Upper Saddle River, NJ: Prentice Hall.

Hutchins, R. (1986). "Having a Hard Time with Just-in-Time." *Fortune*, 64-66.

Imai, M. (1986). *Kaizen, The Key to Japan's Competitive Success*. New York, NY: Random House.

Kawasaki, S. & McMillan, J. (1987). "The Design of Contracts: Evidence from Japanese Subcontracting." *Journal of the Japanese and International Economy*, 1, 327-349.

Klotz, D.E. & Chatterjee, K. (1995). "Dual Sourcing in Repeated Procurement Competitions." *Management Science*, 41(8), 1317-1327.

Kreps, D. & Wilson, R. (1982). "Sequential Equilibria." *Econometrica*, 50, 863-894.

Laffont, J.-J. & Tirole, J. (1986). "Using Cost Observation to Regulate Firms." *Journal of Political Economy*, 614-641.

Laffont, J.-J. & Tirole, J. (1988). "The Dynamics of Incentive Contracts." *Econometrica*, 1153-1175.

Loeb, M.P. & Surysekar, K. (1998). "Payment Ceilings in Cost-Plus Contracting." *Management Accounting Research*, 9, 311-327.

McAfee, R. & McMillan, J. (1986). "Bidding for Contracts: A Principal-Agent Analysis." *RAND Journal of Economics*, 326-338.

McAfee, R. & McMillan, J. (1987). "Competition for Agency Contracts." *RAND Journal of Economics*, 296-307.

Monden, Y. (1983). *The Toyota Production System*. Norcross, GA: Industrial Engineering and Management Press.

Porteus, E.L. (1985). "Investing in Reduced Setups in the EOQ Model." *Management Science*, 31(8), 998-1010.

Reichelstein, S. (1992). "Reliance Investment under Negotiated Transfer Pricing: An Efficiency Result." Working Paper

Riordan, M.H. & Sappington, D.E.M. (1989). "Second Sourcing." *RAND Journal of Economics*, 20(1), 41-58.

Seshadri, S., Chatterjee, K. & Lilien, G.L. (1991). "Multiple Source Procurement Competitions." *Marketing Science*, 10(3), 246-263.

Simchi-Levi, D., Kaminsky, P. & Simchi-Levi, E. (2000). *Designing and Managing a Supply Chain*. New York, NY: McGraw Hill.

Weitzman, M.L. (1980). "The 'Ratchet Principle' and Performance Incentives." *Bell Journal of Economics*, 11(1), 302-308.

Zangwill, W.I. (1987). "From EOQ towards Zero Inventory." *Management Science*, 33, 1209-1223.

THE IMPACT OF EMPLOYMENT AND EXTRACURRICULAR INVOLVEMENT ON UNDERGRADUATES' PERFORMANCE IN A BUSINESS STATISTICS COURSE

Mikhail Kouliavtsev, Stephen F. Austin State University

ABSTRACT

Using data covering every semester and summer from Fall 2007 through Spring 2011 collected at a comprehensive regional public university in an applied business statistics course we examine the impact of academic course load, employment and other time commitments on students' eventual grades in the course. We find that employment has a weak but positive effect on performance in the course, while working more hours lowers one's grade. Involvement in extracurricular activities has no significant effect on grades. We also demonstrate that controlling for students' overall academic ability (we use GPA) is vital to establishing reliable links between these factors and a grade in a particular course.

INTRODUCTION

The purpose of this study is to assess the impact of various academic and non-academic factors on students' academic performance. These include academic course load, on- and off-campus employment, participation in sports and other extra-curricular activities, among others. College students choose to work while enrolled in school in increasingly larger numbers; this is partially due to the rapidly increasing cost of higher education (even in inflation-adjusted terms), but other reasons have been cited in the literature as well.[1] And there are other activities competing for the student's time (and attention): participation in sports – whether formal or informal (e.g., intramurals), involvement in extracurricular activities – such as social clubs, service and Greek organizations; volunteer work, etc. The central question of interest here is: Does greater involvement in these non-academic activities adversely affect academic performance?

Briefly, two competing theories can be advanced with respect to the likely direction of this effect. First, outside activities take up time, at least some of which would otherwise be devoted to schoolwork, which in turn can lead to lower grades. Second, students choosing to work or get involved in clubs and organizations are better at prioritizing work, managing their time, and may in general be more motivated than an average student; if so, then this selection bias would lead us to finding students with greater involvement also performing better

academically. It is likely, however, that both effects are present to some degree (as previous studies have documented – see below): for example, working several hours a week is associated with receiving better grades, but working too much can lead to poorer performance.[2]

To address these issues (and others), we employ a unique dataset assembled over five years in an applied business statistics course taught in a AACSB-accredited college of business at a US comprehensive regional public university.

LITERATURE REVIEW

Studies linking college students' employment to their academic performance are rather few and far between. Even fewer studies consider the effects of other factors – such as extracurricular involvement – on academic performance. A notable exception is Dundes and Marx (2007) who study academic performance (defined as "hours studied and higher grades") of undergraduate students at a small liberal arts college, and find that students who strike the optimal balance between work and school perform the best. That balance turns out to be between 10 and 19 hours of work per week; moreover, students in this group do better academically on average than all other students, including those who do not work at all.

These findings are echoed by Pike et al. (2008) who use data from the National Survey of Student Engagement (NSSE) to examine the effect of work on first year students' academic achievement. They conclude that working more than 20 hours a week affects grades negatively, even after controlling for a variety of individual student characteristics. The study also suggests that working on campus for 20 hours or less has an indirect positive effect on first year students' grades acting through increased engagement. Similarly, Torres et al. (2011) find that younger undergraduates at two urban commuter colleges in Indiana who work more than 31 hours a week have lower grade point averages (GPAs) and are less likely to complete the courses they enroll in.

A different, more theoretic modeling approach is taken by Wenz and Yu (2010), who actually model a student's decision to work while enrolled in school. In general, they find that working has a negative impact on academic achievement; specifically, each work-hour lowers student GPA by 0.007. More interestingly though, the authors conclude that there are significant differences in the academic achievement of various groups of students depending on what motivates them to work in the first place: students working for primarily financial reasons earn lower grades than those who work to acquire career-specific skills, but higher grades than students who simply want general work experience.

A contrasting set of results is obtained by Bradley (2006) who reports that the study of 246 university students found approximately 85 percent reporting having a paid job during the semester. However, academic performance appeared to be unaffected by either employment or the number of weekly hours worked: GPA's were relatively high for both, those who did not work and those working more than 20 hours per week.

A number of studies report similar results from analyses of working college students in other countries. For example, Humphrey (2006) uses data obtained from a survey of students at an English university to establish a significant link between working and end-of-year average grade as well as participation in university societies. A key finding is that students coming to the university from state schools tend to work more than students from private schools, and as a consequence tend to, on average, perform worse academically.

Callender (2008), perhaps the most comprehensive study in its scope – the data used are from 1,000 students at six British universities – also finds that working has a detrimental effect on the students' final year marks as well as their degree results (i.e., graduation). In particular, students working an average number of hours per week are "a third less likely to get a good degree than an identical non-working student."

A study of Chinese college students by Wang et al. (2010) finds that working part-time has no effect on academic performance of students and a positive effect on the students' social life, while possibly damaging their relationship with their parents.

There are even fewer examples in the literature of studies addressing the effects of other student activities on academic achievement and performance. Turley and Wodtke (2010) report that while in general first-year college students living off-campus perform as well as those living on-campus, among Black students, campus residents have higher GPA's than similar students at the same institution living off-campus. Rees and Sabia (2010) find some limited support for the theory that sports participation increases motivation, teaches teamwork and self-discipline, thereby leading to better performance in school.

Quite a few studies report results of data from high school students engaging in extracurricular activities and paid work (Lee & Orazem, 2010; Staff et al., 2010; Patton & Smith, 2010). Many of the studies conclude that working more hours during the academic year does not affect high school academic performance; on the other hand, increased high school work intensity raises the probability of completing high school but lowers the likelihood of going to college. In any event, this lies outside of the scope of our paper.

While the existing literature does have something to say about the link between student employment and academic performance, our study improves and advances our understanding of the subject in several ways. First, as discussed above, many existing studies focus on high school students' academic performance rather than college students; this is mainly due to lack of data on undergraduates' experiences. Second, most studies measure academic performance using standardized test scores, whereas we use grades actually earned in a specific course, which is clearly a superior measure of performance *in that course*. In particular, our data are from the same course taught by the same instructor using the same approach (e.g., a consistent grading scheme), all of which makes it easier to isolate the effects of other factors on student performance. Third, the present study is the first, to the authors' knowledge, to focus on a business statistics course. This is important because this particular subject is unique in the sense that the material taught combines ideas, analytical thinking, ability to abstract, quantitative skills,

and problem solving – in other words, many of the knowledge areas that college students are expected to acquire. Fourth, the data we use are substantial in their depth and breadth, containing detailed information at the individual student level spanning hundreds of students, several years and settings (e.g., morning vs. afternoon or evening class), all of which make this vastly superior to using data from a large but generic dataset, such as the National Longitudinal Survey of Youth (NLSY).

BACKGROUND

The course, "Applied Statistical Analysis," ("statistics" hereafter) is part of the "business foundation" and is required for all business majors.[3] The course is typically taken by students in their junior year and has one prerequisite – Probability and Statistics. This prerequisite course is taught in the Mathematics department, outside of the College of Business. Many students take this prerequisite elsewhere, most often at a community college.[4]

Statistics is a basic business statistics course covering topics of hypothesis testing, ANOVA, and regression techniques. Other topics – such as nonparametrics – are sometimes covered by individual faculty teaching the course but are not required. The focus is on applied data analysis with heavy emphasis on business and economic applications.

Students' grades in this course, when taught by the author, are determined as follows. There are four noncumulative exams and eight graded quizzes. The quizzes are administered online through a course management system (such as Blackboard) and contribute 15 percent to the overall grade. The exams are all equally weighted at 20 percent each. An additional five percent of the grade is reserved for class attendance and participation.

Students are given an opportunity to improve their grade by completing an optional project using regression techniques they learned in class. The project involves the student independently selecting a topic to study, formulating a hypothesis, deciding what dependent variable and independent variables to use in testing the hypothesis, locating or collecting data, and performing regression analysis. The topic can be anything of the student's choosing, but must be approved by the instructor on the criterion of being "doable" – i.e., variables are quantitatively measurable, data are available or can be reasonably gathered, etc. The grade for the project, which comes from a presentation the student makes to the class and a paper submitted to the instructor, replaces the lowest of the first three exam grades; students cannot use the project to avoid taking the fourth exam.

DATA SOURCES AND CONSTRUCTION

In this section, we briefly describe how the data were obtained and the methodology of analysis used. On the first day of every semester and summer session from fall of 2007 through

spring of 2011, the author asked students to fill out a 3" X 5" card with the following information:

Name
Phone number
Email address
Major
Year in school
List of mathematics courses student has taken previously
How many credit hours student is taking during the current term
Other time commitments student has

The phone number and email address are collected for record-keeping purposes only. Asking for this information was more important before it was easily obtainable through a university information portal. Verbally, the students are encouraged to list any math courses they have taken in college -- i.e. to not include any high school courses or advanced placement credit, but to include courses transferred from, say, a junior college. The last two items are the most interesting since the first provides us with the course load the student has at the time she is taking the course, while the second reveals what other activities she is involved in. This is where students tell us if they work (and if so, how many hours per week), participate in sports or other activities (again, with an estimate of a weekly time devoted to these activities) or anything else they deem important and time consuming. Examples of what students have listed in the past include: Taking care of family (children, elderly or disabled relatives), informal sports (e.g., fitness), church and other faith-based activities. The remaining items on this survey provide sources of control variables – student's major, year in school, level of math background/preparation, and when the class was taken (semester, year, day of the week, and time of day).

These data are matched with student records (obtained from the university registrar's office) on cumulative GPA and author's own records containing students' final course averages.

Our dependent variable – the variable of interest – is of course the student's grade, and it is modeled as a function of the other (explanatory) variables. We pose and attempt to answer the following questions:

- Does working (i.e., having a job) significantly affect students' grades?
- Is there a significant difference between working on- and off-campus in terms of its effect on grades?
- Does working more hours per week have a significant negative effect? If so, does it become more pronounced above a certain number of hours per week?

- Does involvement in extracurricular activities have an effect on grades?
- Is an increase of one hour per week in an extracurricular activity equivalent to working an additional hour at a job in terms of its effect on grades?
- What is the effect of a heavier academic course load on grade in a given course?
- Does the level of math preparation or background matter for one's grade in statistics?
- Are courses taken in the summer different from those taken in the long semester or at certain times of day, holding other factors constant?
- Are any of the above effects stronger for certain majors?

SUMMARY OF DATA

The dataset is comprised of 554 observations. This covers the period from Fall 2007 through Spring of 2011 and includes classes taught during summer sessions in addition to those taught during long semesters. Only students who completed the course were retained in the sample – i.e., those who stopped coming to class or failed to take all four of the scheduled exams, were discarded.[5] Students taking the course multiple times are retained as multiple independent observations, provided they completed the course each time. Below we highlight some summary statistics, which are shown in Table 1.

Table 1: Descriptive Statistics

Variable	Mean	Std. Dev.	Min	Max
grade	81.609	10.295	45.74445	100
gpa	2.831	0.565	1	4
course_load	12.903	3.714	3	20
pct_full_load	0.913	0.192	0.4	2
math_courses	2.638	0.965	1	7
spring	0.379	0.486	0	1
fall	0.507	0.500	0	1
summer	0.114	0.318	0	1
am	0.841	0.366	0	1
before10	0.446	0.498	0	1
accounting	0.181	0.385	0	1
management	0.146	0.354	0	1
finance	0.094	0.292	0	1
marketing	0.208	0.406	0	1
economics	0.022	0.146	0	1
generalbus	0.233	0.423	0	1
intlbus	0.054	0.227	0	1
compsci	0.009	0.095	0	1
other	0.023	0.152	0	1
two_major	0.029	0.168	0	1

Table 1: Descriptive Statistics				
Variable	**Mean**	**Std. Dev.**	**Min**	**Max**
remedial	0.038	0.191	0	1
advanced	0.031	0.173	0	1
freshman	0.005	0.073	0	1
sophomore	0.027	0.162	0	1
junior	0.417	0.494	0	1
senior	0.534	0.499	0	1
5th_year	0.005	0.073	0	1
work	0.610	0.488	0	1
work_hours	15.627	15.022	0	60
on_campus	0.058	0.234	0	1
activities	0.572	0.797	0	5
sports	0.081	0.273	0	1
greek	0.119	0.324	0	1
volunteer	0.022	0.146	0	1
family	0.043	0.204	0	1
church	0.020	0.140	0	1

About 38 percent of the sample took the course in the Spring; about 51 percent took it in the Fall, with the remaining students taking it in the Summer (second five-week session of the summer term, to be exact.) Approximately 84 percent of the sample took the course when it was scheduled to start in the morning (before noon); about 45 percent took the course with a pre-10:00 am scheduled start.

The mean numeric course grade (*grade*) for the entire sample is 81.58 with the range 45.75 to 100. As noted before, this final course grade reflects some students' attempts to improve their performance by completing the optional data analysis project described earlier; therefore, this mean likely overstates the students' average performance in the class (as reflected by exams and quizzes only).

The majors represented in the sample are as follows:

General Business	23%
Marketing	20
Accounting	18
Management	15
Finance	9
International Business	5
Economics	2
Computer Science	1
Other	2
Total	100%

In addition, about 2.9 percent of the sample reported having a double major. This is fairly representative of the relative popularity of the business majors in general, which should come as no surprise since all business majors must take the statistics course.

The mean reported number of math courses taken prior to enrolling in statistics is 2.63, with the range from 1 to 7. A minimum of one course is reported for all observations, which is re-assuring since there is a one-course pre-requisite. About 3.8 percent of the sampled students reported having taken at least one remedial math course, while about 3.2 percent reported having taken "advanced" math courses.[6]

Overwhelmingly, the students in the sample are self-reported juniors and seniors, with the exact breakdown of classifications as follows:

Freshman	0.6%
Sophomore	2.7
Junior	42
Senior	53
5th year	0.5
Total	100%

Mean course load is 13.88 semester credit hours for the long semester (ranging from 6 to 20 hours) and 4.66 for the summer session (ranging from 3 to 12, though the 12 credit hours reported during the summer term is an extreme outlier and almost certainly includes courses taken at another institution and/or online concurrently with courses taken at this university.) To combine long semester and summer observations into a single measure of a student's course load, we compute the percentage of a full course load (*pct_full_load*), taken to be 15 hours in Spring/Fall and 6 hours in Summer, for each observation. The mean percentage of a full course load for the entire sample is 91.2. In other words, most students are either full-time college students, in the traditional sense, or close to it.

About 61 percent of the sample report having some type of gainful employment while in school. The overall average number of hours worked per week is 15.6; the average number of hours worked for just those who reported working is 25.22, ranging from 4 to an unbelievable 60 hours per week! Nearly 10 percent of those who reported working (approximately 5.7 percent of the sample) said they worked on campus; in some instances these students worked off-campus as well.

In response to the question prompting the students to list "other significant time commitments," the mean number of extra-curricular activities reported is 0.6, ranging from zero to 5. The mean number of activities for just those who reported such activities is 1.36. Approximately 8.1 percent reported being involved in sports; slightly less than 12 percent were involved in a Greek organization (social fraternity or sorority); about 4.3 percent cited taking care of family as an activity, which included children, parents, as well as extended family; 2 percent reported church involvement, and 2 percent reported volunteering.

To control for the variation across students in academic ability, we use cumulative grade-point average (GPA) at the time the student enrolls in statistics. The observed values of cumulative GPA in the sample range from 1.0 to 4.0 with a mean of 2.83 and sample standard

deviation of 0.57. We expect the students' GPAs to account for a large portion of the variation in course grades.[7]

RESULTS

We estimated several models on our complete dataset. The results are shown in Table 2. Some of the factors appear to be significant predictors of grades in every specification that we attempted, so we discuss those here first, before turning to other variables of interest.

Table 2: Estimation Results
(absolute value *t*-statistics in parentheses)

Variable	Model 1	Model 2	Model 3	Model 4	Model 5
Constant	48.614***	51.726***	52.243***	50.211***	85.007***
	(28.51)	(27.97)	(26.30	(21.00)	(49.14)
GPA	12.28***	12.287***	12.096***	11.608***	
	(21.95)	(22.34)	(21.49)	(19.81)	
Spring		-2.302**	-2.345**	-2.893**	-2.038
		(2.11)	(2.12)	(2.54)	(1.35)
Fall		-4.415***	-4.389***	-4.912***	-4.591***
		(4.19)	(4.13)	(4.44)	(3.17)
Before10	-3.413***	-2.781***	-2.847***	-2.81***	-2.473***
	(5.44)	(4.26)	(4.34)	(4.24)	(2.76)
Work	2.104*				
	(1.84)				
Work_Hours	-0.0982***	-0.0351*	-0.041*	-0.0243	-0.104***
	(2.63)	(1.70)	(1.90)	(1.09)	(3.57)
On_Campus			1.662	1.922	3.2*
			(1.24)	(1.42)	(1.76)
Math_Courses			0.0785	0.0877	0.901**
			(0.24)	(0.26)	(1.99)
Remedial			-2.607	-2.115	-9.005***
			(1.54)	(1.25)	(3.96)
Pct_Full_Load				3.049*	
				(1.75)	
Activities			-0.0975	-0.165	0.435
			(0.24)	(0.33)	(0.79)
Sports				0.893	
				(0.71)	
Greek				0.361	
				(0.33)	
Volunteer				-0.07	
				(0.03)	
Family				1.581	
				(1.00)	
Church				-0.283	
				(0.12)	
Accounting				1.782**	
				(2.12)	

Table 2: Estimation Results (absolute value *t*-statistics in parentheses)					
Variable	Model 1	Model 2	Model 3	Model 4	Model 5
Economics				0.695	
				(0.32)	
Finance				2.856***	
				(2.61)	
Compsci				2.89	
				(0.88)	
Adj. R^2	0.493	0.509	0.509	0.513	0.094

Cumulative GPA (*gpa*) is by far the most important factor – both in terms of its statistical significance and in terms of the magnitude of its impact – in predicting the course grade. In all of our specifications, *gpa* is highly significant, and an increase of one point in this variable is associated with an increase in *grade* of between 11.6 and 12.3 percent, i.e. more than a full letter grade. It appears that students who took statistics in the fall did significantly worse than students who took the class during the spring or summer sessions. In particular, taking the class in the fall lowers one's grade by between 4 and 5 percentage points relative to taking the class in the summer, depending on the exact model specification. The coefficient on the spring semester dummy is also significant and negative in all of our regressions but has a smaller magnitude, averaging about 2.3 percent, suggesting a negative effect there as well. The impact of the long-semester classes is likely underestimated here: summer classes are too short to allow for the optional project opportunity, so the grades students receive in the summer, everything else the same, should on average be lower. This observation is difficult to explain. Perhaps summer courses allow students to concentrate their efforts better; also, it is possible that a class that meets every day reinforces students' learning in a way that a class that meets only twice a week simply cannot. It is worth noting that it does not seem as though higher grades in the summer are a result of students not taking a multitude of other courses (our measure of academic course load is only very weakly significant in one of our specifications, where it actually has a positive coefficient) or working substantially less (the average number of hours worked per week in the summer is 23.8 versus 25.2 during the academic year).

Students taking the class in the early morning (before 10 a.m.) receive between 2.5 and 3.4 percent lower grades than those taking later classes. This is not entirely unexpected. Anecdotally, a typical college student prefers afternoon classes to morning classes. This is supported by the author's own observation that whenever sections of the same course are scheduled both in the morning and in the afternoon, the afternoon sections tend to fill up with registered students first. While this does not automatically imply higher grades in the afternoon courses, it is likely that the relatively stronger (academically) students are first to sign up, while the weaker students tend to postpone registering. Furthermore, students needing to repeat the course due to failing it during the previous term would be registering late and likely forced in to the remaining open spots in the morning classes.

We now turn to the two other main factors of interest – work and involvement in extracurricular activities. The coefficient on the *work* dummy is positive and significant (but only at the 10 percent level), indicating that students who work actually do better in statistics than those who do not, on average. On the other hand, the coefficient on *work_hrs* is significant and negative, suggesting, as expected, that working more hours per week tends to lead to lower grades. Specifically, in our Model 1, while working is associated with about a 2.1 percent increase in the course grade, working each additional hour per week lowers one's grade by 0.1 percent. The latter effect is obviously too small to be meaningful, even if one considers a large increase in weekly hours worked: say, going from 20 to 40 hours per week. Furthermore, this effect does not appear robust as the significance of *work_hours* drops and eventually goes away completely as we add more factors to the model.

What is perhaps most interesting (and somewhat surprising) is that we find no evidence of other, non-work activities having any effect on student grades. None of our dummy variables turn out to be significant predictors of grades. The sheer number of activities reported as "significant time commitments," recorded in *activities* comes out insignificant as well. It is possible that some of the activities go underreported. For example, if a student is not a member of a social fraternity or club, he may not report being involved at all, whereas in reality the formal activity "Greek fraternity" may simply be substituted by the informal activity "hanging out with friends". The latter is still a time-consuming activity and could still have an effect on one's grades, but our data have no record of it.

The remaining few control variables are included to determine whether performance varies across student classifications or majors. We find some evidence of differences across majors in Model 4: students majoring in finance receive significantly higher grades – by almost 3 percent – while accounting majors enjoy a nearly 2 percent boost relative to other majors. This result makes some intuitive sense if one considers that finance and, to a lesser extent, accounting, are relatively more quantitative fields than, say, marketing; in other words, it is possible that students majoring in these areas tend to be better performers in other quantitative courses outside of their majors.

We attempted a "kitchen sink" kind of a regression model as well – adding every regressor available, including full sets of dummies for all majors and student classifications, to the right hand side of our model. The results did not reveal anything not already discussed above, so out of space considerations they are relegated to an appendix, available from the authors upon request.

As one final exercise, we removed *gpa* from our model and estimated a "naïve" regression – essentially, Model 3 but without the GPA variable; the results are shown as Model 5 in Table 2. Not surprisingly, the fit of the model as measured by R^2 drops considerably from about 0.51 for Model 3 to only 0.094 for Model 5. More importantly, however, several factors emerge as significant predictors of grades whereas previously (i.e., in a model controlling for students' GPA) they were not.

Of particular interest is the impact of mathematical background and preparation on student performance in our statistics course. Students who reported taking at least one remedial math course did significantly worse in statistics: the coefficient on the *remedial* dummy is about –9. This is a rather large effect, amounting to as much as a letter grade. On the other hand, the number of math courses taken previously has a positive effect on grades, albeit the coefficient is small – less than 1 – but is statistically significant at the 5 percent level. Clearly, these results are biased – we showed above that once student academic quality is controlled for, neither of the math preparation measures matters. In fact, it appears that strong students find a way to do well in the course regardless of their mathematical background. However, we present these results here as a cautionary note: it would be rather easy to draw conclusions based on the results in Model 5, which suffer from the omitted variable bias.

We would be remiss to not at least mention one important caveat plaguing our analysis. The decision on the part of a student to work and/or get involved in activities (and for that matter, how many hours to work or how many activities to participate in) is endogenous in our model. In other words, a student may decide to reduce the number of hours she works after noticing her grades declining as the semester progresses; alternatively, a student may choose to get a job midway through the term once he realizes that he has a firm grasp on his studies and has some free time. Similarly, students can drop courses to lighten their load part way through the semester. Unfortunately, we are unable to track such changes for each student; our data collection occurs at the very beginning of the term, so we implicitly assume that the student has made whatever choices she will make with respect to her course load, employment and extracurriculars. On the other hand, we have no reason to suspect that students who, say reduce their work hours or participation in activities outnumber those who increase their workload or get more involved. Therefore, we do not suspect that our results are systematically biased in any direction.

CONCLUDING REMARKS

In this study of student performance in a typical business statistics course, we find that a student's overall academic ability – proxied by GPA – is the single most important factor. When students take the course also has some effect (e.g., summers are better than either spring or fall semesters, and afternoons are better than mornings). On the other hand, being involved in many extracurricular activities has no measurable impact on grades; having a job has a weak positive effect on grades, while the effect of working more hours has a small but significant negative effect. We also demonstrate the risk of drawing conclusions based on a "naïve" model which omits a key determinant of student success in a particular course – her overall GPA.

Our results should be of interest to a wide audience: current and future college students, parents, faculty, and academic advisors. One potential way to extend and improve our approach and results would be to consider how certain components of a student's grade, such as the quiz

average or attendance, are affected by working or being involved in many activities. For example, it would be interesting to examine how important class attendance is to performing well in the course, and if those with many other time commitments are able to keep up with regular assignments (quizzes).

ENDNOTES

[1] Other reasons commonly cited for a college student's decision to work while in school: financial stress (unrelated to the cost of college), need to obtain experience for a future job, and boredom.

[2] There is also a third possibility: students who work may actually become better students over time as working helps them develop time- and task-management skills.

[3] The following undergraduate business majors are offered within the College of Business: accounting, business economics, computer information systems, finance, general business, international business, management, and marketing. In addition, nonbusiness majors are offered in computer science (BS), economics (BA), and information technology (BA). Of these, only economics majors are required to take statistics.

[4] About 29 percent of all students taking statistics during the period covered by our data took the pre-requisite course elsewhere; most of those (about 27 percent of our sample) took it at a junior college.

[5] This does not include students who dropped the course at some point during the semester; those students do not receive a letter grade at all.

[6] "Advanced" math is (somewhat arbitrarily) defined here to include courses such as Calculus II and beyond. Basically, any math courses outside of the typical math sequence required for Business majors (or equivalent course substitutions) reported by students are considered advanced.

[7] The GPA we use may not be a very good measure of a student's actual collegial academic performance because many students transfer courses to the university from other colleges, most often two-year junior colleges. These transferred credits do not impact the student's GPA calculation, which only takes into account courses completed in residence. While we do not suspect that this biases the GPA measure (on average) either upward or downward, it is worth keeping in mind that for some students, as many as 60 credit hours completed at another institution could, theoretically, be "off the record" vis-à-vis GPA calculation.

REFERENCES

Bradley, Graham. 2006. Work Participation and Academic Performance: A Test of Alternative Propositions. *Journal of Education and Work,* 19(5), 481-501.

Callender, Claire. 2008. The Impact of Term-Time Employment on Higher Education Students' Academic Attainment and Achievement. *Journal of Education Policy*, 23(4), 359-77.

Dundes, Lauren and Jeff Marx. 2006-07. Balancing Work and Academics in College: Why Do Students Working 10 to 19 Hours Per Week Excel? *Journal of College Student Retention: Research, Theory & Practice,* 8(1), 107-20.

Humphrey, Robin. 2006. Pulling Structured Inequality into Higher Education: The Impact of Part-Time Working on English University Students. *Higher Education Quarterly*, 60(3), 270-86.

Lee, Chanyoung and Peter F. Orazem. 2010. High School Employment, School Performance, and College Entry. *Economics of Education Review*, 29(1), 29-39.

Patton, Wendy and Erica Smith. 2010. Part-Time Work of High School Students: Impact on Employability, Employment Outcomes and Career Development. *Australian Journal of Career Development*, 19(1), 54-62.

Pike, Gary R., George Kuh and Ryan C. Massa-Mckinley. 2008. First-Year Students' Employment, Engagement, and Academic Achievement: Untangling the Relationship Between Work and Grades. *NASPA Journal*, 45(4), 560-82.

Rees, Daniel I. and Joseph J. Sabia. 2010. Sports Participation and Academic Performance: Evidence from the National Longitudinal Study of Adolescent Health. *Economics of Education Review,* 29(5), 751-59.

Staff, Jeremy, John E. Schulenber and Jerald G. Bachman. 2010. Adolescent Work Intensity, School Performance and Academic Engagement. *Sociology of Education*, 83(3), 183-200.

Torres, Vasti, Jacob P. K. Gross and Afet Dadashova. 2010-11. Traditional-Age Students Becoming At-Risk: Does Working Threaten College Students' Academic Success? *Journal of College Student Retention: Research, Theory & Practice,* 12(1), 51-68.

Turley, Ruth N. Lopez and Geoffrey Wodtke. 2010. College Residence and Academic Performance: Who Benefits from Living on Campus? *Urban Education*, 45(4), 506-32.

Wang, Hongyu, Miosi Kong, Wenjing Shan and Sou Kuan Vong. 2010. The Effects of Doing Part-Time Jobs on College Student Academic Performance and Social Life in a Chinese Society. *Journal of Education and Work*, 23(1), 79-94.

Wenz, Michael and Wei-Choun Yu. 2010. Term-Time Employment and the Academic Performance of Undergraduates. *Journal of Education Finance*, 35(4), 358-73.

EXAMINING NCAA/NFL MARKET EFFICIENCY

Gerald Kohers, Sam Houston State University
Mark Tuttle, Sam Houston State University
Donald Bumpass, Sam Houston State University

ABSTRACT

Billions of dollars are wagered every football season in the hopes of identifying a winning strategy. Numerous articles have been conducted that try and identify inefficiencies in football betting markets. While many studies have determined that the betting market is efficient, there are a few articles that have identified potential anomalies in the football betting markets. The sustainability of these anomalies is still another question. This study takes an alternative approach in testing the market efficiency of betting on NCAA college football and National Football League (NFL) games. A unique, primary source data set was collected between 2003 and 2011, which offers some additional insight in the study of market efficiency that has not been included in previous studies. The results of this study indicate that the betting market is indeed efficient.

INTRODUCTION

Gambling on the outcome of an event dates back to the Chinese, Japanese, Greeks and Romans in 2300 B.C. In the United States, gambling dates back to the Native Americans and early Colonists (Roberts, 1997). From cockfights, horse races, and bare-knuckled fights, Americans throughout history have been drawn towards picking a winner (Martin, 2012). In the U.S., Gambling has evolved into a multi-billion dollar industry with significant economic impacts (Bazelon, Neels, & Pallavi, 2012). NCAA college football and the National Football League (NFL) are two of the most popular forms of gambling (IntenseGambling.com, 2012).

There are a variety of different forms of betting on football, the most popular of which is using a point spread. In this form of betting, the underdog team is "spotted" the point spread amount. In order for the favorite team to win the bet, their score must exceed the point spread differential. Thus the point spread that the bookmakers come up with has a huge impact and miscalculations may result in major gains or losses. Bettors are constantly trying to identify opportunities to win. However, if the football betting market is efficient then there would not be any sustainable opportunities.

The bookmakers also receive their "take" of the bet, called "the vig." A standard football bet would be to place a $110 wager in order to win $100. If the bettor is correct, then they would receive $100 along with $110 that they initially bet. Factoring the vig into the bet means that in

order for bettors to realize a profit, they need to be correct more than 52.38% (110/210) of the time.

Numerous studies have been conducted testing the efficiency of the NFL betting market. A number of studies find the NFL betting market to be efficient (Boulier, Steckler, & Amundson 2006; Dare & Holland, 2004; Boulier & Steckler, 2003). There have been other studies that have identified possible betting strategies that may lead to potential profits. For example Borghesi, 2007a revealed that home teams and the weather could be taken advantage of. Borghesi, 2007b identified late season biases and Paul & Weinbach, 2002 noticed an over-under betting strategy that would yield small profits, although they did not believe this anomaly would last over a long period of time.

The college football market differs from the heavily watched professional market. College coaches do not have to reveal injury reports, and players on professional teams are very well known, whereas many freshman/sophomore players on college teams are not. This lack of information may provide local bettors with additional information that the bookmakers and non-local bettors don't have easy access to.

Efficiency in the NCAA football betting market has not been tested as frequently as the NFL betting market, yet has similar results. A number of these studies have shown the efficiency of the NCAA Football betting market (Fair & Oster, 2005; Paul *et. al,* 2003; Dare & McDonald, 1996; Golec & Tamarkin, 1991). Sinkey & Logan, 2009 however, point out that many of the previous studies used limited data and used indirect tests of market efficiency. In their study, they find that the NCAA market is inefficient and profits can be made by looking at a combination of home teams, favored teams with strong traditions and on teams that play weak opponents. They find that the inefficiency is due to the bookmakers response to the gamblers behavioral bias.

DATA/METHODOLOGY

Given the mixed results from the previous studies, this research examines market efficiency in football betting markets. Specifically, the data in this study is from local bettors in a small geographical area. In order to test the efficiency of football betting markets, a unique approach for data collection was undertaken. For a given week during the football season (September – December) ten games were selected, five college games and five NFL games. The five college games were selected based on a few factors, national ranking, regional interest, and whether the game was televised. The five NFL games were selected similarly. Typical games would be those that included the Houston Texans, the Dallas Cowboys, Sunday and Monday night games, teams with good records and were nationally ranked and games that were locally televised.

Participants were asked to pick the favorite or the underdog for the given point spread. Point spreads were chosen on Wednesday mornings from an aggregate collection of published

point spreads. They were to rank the certainty of each game from 10 to 1 (using each number once). For a given point spread, a 10 was assigned to the team that the individual thought would most certainly cover or beat the spread, and a 1 to the individual's least certain pick. If their selection was correct they would receive those points. The winner for the week was the individual with the most points. Thus, if the individual was right on all 10 games they would get the maximum of 55 points (10 + 9 + … + 1). As an added incentive, individuals would pay $5.00 each week to play with the winner receiving the entire entry fee, i.e. winner takes all. The forms were due in before the first game on the sheet, which was typically Saturday mornings. A sample entry sheet is shown in Figure 1 below.

The data was collected from 2003 – 2011. Each year the number of weeks ranged from 12 to 15. The discrepancy in numbers is a result of whether or not Thanksgiving week was included, if a natural disaster occurred, or the inability to get the picks out. This resulted in 123 weeks of picks. On average, ~34 individuals would pick each week, with the most number of picks in one week being 52 and the least number of picks being 10. The total number of picks were 4,074.

RESULTS

In order to test the market efficiency of the football gambling market one would hypothesize that if the market is efficient, then the average score of the picks would be 27.5 (55 possible points/2).

H_0 *μ = 27.5. If the football betting market is efficient, the average score of the picks will not deviate from the overall mean of 27.5.*

The average from our data resulted in an average score of 28.7. Performing a t-test on the mean indicates the average score is significantly different from 27.5 with a p-value of 0.00. Thus, given this unique data sample, the results suggest that the NCAA and NFL betting markets may not be efficient. However, factoring the vig in, then the average score that the bettor must obtain is 28.81 (55 x .5238). Testing this hypothesis:

H_0 *μ = 28.81. If the vig is factored into whether the football betting market is efficient, the average score of the picks will not deviate from the mean of 28.81.*

The results of this test indicate no significant difference exists between the means at a p-value of 0.46.

Figure 1: Sample Data Sheet

Name: _____ **Phone:** _____

FOOTBALL PICKS Week 4 e-mail: _____

$5.00 Entry

Pick the winner of each game with the point spread.
Rank the certainty of each game from 10 to 1 (use each number once).
10 is most certain, 1 is least certain.
Put the # next to the team you think will win with the point spread.
The winner is the person with the most points.
Home Team is in ALL CAPS

	FAVORITES		**UNDERDOGS**	
	College Games	Points		
_____	Texas A&M (14) (11:00)	3	ARKANSAS (18)	_____
_____	SO CAROLINA (10) (2:30)	9 1/2	Auburn	_____
_____	VA TECH (11) (5:00)	7	Clemson (13)	_____
_____	Alabama (3) (7:00)	3 1/2	FLORIDA (12)	_____
_____	WISCONSIN (7) (7:00)	9	Nebraska (8)	_____
	NFL Games	Points		
_____	DALLAS (12:00)	1 1/2	Detroit	_____
_____	HOUSTON (12:00)	4	Pittsburgh	_____
_____	NY Giants (3:05)	1 1/2	ARIZONA	_____
_____	BALTIMORE (7:25)	3 1/2	NY Jets	_____
_____	TAMPA BAY (7:35)	10	Indianapolis	_____

One could also surmise that if the market were inefficient then the individual's most certain pick, their 10 pointer, would be correct more often than their least certain pick, their 1 pointer. Table 1 provides the number of correct/incorrect picks for each certainty level. There is

very little variation in the picks. However, in all cases the bettors were correct more than half (2,037) the time.

Table 1: Number of Correct/Incorrect Picks by Certainty Level (n = 4,074)		
Certainty Level	Correct	Incorrect
10	2,185	1,889
9	2,131	1,943
8	2,148	1,926
7	2,072	2,002
6	2,068	2,006
5	2,131	1,943
4	2,133	1,941
3	2,099	1,975
2	2,091	1,983
1	2,118	1,956

Table 2 contains the p-values from the tests of equality between the average number of correct certainty picks compared to each of the other correct certainty picks. The results do not indicate that the individuals' most certain pick is significantly better than their least certain pick. For example, the average number of times individuals were correct with their certainty score 1 selection is statistically equal to the number of times individuals were correct with their certainty score 10 selection (p-value of 0.48).

Table 2: Equality of Means Test Results (p-values)									
Certainty Score	Certainty Score								
	2	3	4	5	6	7	8	9	10
1	0.77	0.84	0.87	0.89	0.59	0.62	0.75	0.89	0.48
2		0.93	0.65	0.67	0.80	0.84	0.54	0.67	0.32
3			0.72	0.73	0.74	0.77	0.60	0.73	0.36
4				0.98	0.49	0.51	0.87	0.98	0.59
5					0.50	0.53	0.86	1.00	0.57
6						0.97	0.39	0.50	0.21
7							0.42	0.53	0.23
8								0.86	0.70
9									0.57

CONCLUSION

This study looked into the market efficiency of the football betting markets, investigating both the NCAA and NFL markets. A unique data set was collected from local bettors capturing their level of certainty in identifying potential wins in college and professional games. The

results indicate that bettors are correct more than 50% of the time. However, from a profitability stand point, after factoring in the bookmakers take, the vig, the bettors are not better than average. The latter results support the findings that the football betting markets are, in general, efficient.

REFERENCES

Bazelon, Coleman, K. Neels, & P. Seth (2012). Beyond the Casino Floor: Economic Impacts of the Commercial Casino Industry. *American Gaming Association*, February 2012. Retrieved 3/3/2012, from http://www.americangaming.org/files/aga/uploads/docs/final_final_brattle_study_2-3-12.pdf

Borghesi, Richard (2007a). The Home Team Weather Advantage and Biases in the NFL Betting Market. *Journal of Economics and Business* 59: 340-354.

Borghesi, Richard (2007b). The Late-Season Bias: Explaining the NFL's Home-Underdog Effect. *Applied Economics* 39: 1889-1903.

Boulier, Bryan & H.O. Steckler (2003). Predicting the Outcomes of National Football League Games. *International Journal of Forecasting* 19: 257-270.

Boulier, Bryan, H.O.Steckler, & S. Amundson (2006). Testing the Efficiency of the National Football League Betting Market. *Applied Economics* 38: 279-284.

Dare, William & S. Holland (2004). Efficiency in the NFL Betting Market: Modifying and Consolidating Research Methods. *Applied Economics* 36: 9-15.

Dare, William & S. McDonald (1996). A Generalized Model for Testing the Home and Favorite Team Advantage in Point Spread Markets, *Journal of Financial Economics*, 40: 295-318.

Fair, Ray & J. Oster (2005). College Football Rankings and Market Efficiency. Discussion Paper No. 1381. Cowles Foundation for Research in Economics, Yale University, http://cowles.econ.yale.edu/.

Golic, Joseph & M. Tamarkin (1991). The Degree of Inefficiency in the Football Betting Market. *Journal of Financial Economics*, 30: 311-323.

IntenseGambling.com (2012). NCAA & NFL Football Betting. Retrieved 2/5/2012, from http://www.intensegambling.com/sports-betting/online/american-football/.

Martin, Jerry (2012). History of Sports Betting Retrieved 1/22/2012, from http://www.ultimatecapper.com/history-of-sports-betting.htm

Paul, Rodney & A. Weinbach (2002). Market Efficiency and a Profitable Betting Rule: Evidence from Totals on Professional Football. *Journal of Sports Economics* 3: 256-263.

Paul, Rodney, A. Weinbach, & C. Weinbach (2003). Fair Bets and Profitability in College Football Gambling. *Journal of Economics and Finance,* 27(2): 236-242.

Roberts, Mike (1997). The National Gambling Debate: Two Defining Issues. Whittier Law Review, Scarne's New Complete Guide to Gambling; *Encarta Online Encyclopedia*, v8, n3.

Sinkey, Michael & T. Logan 2009. Betting Markets and Market Efficiency: Evidence from College Football. Retrieved 1/22/2012, from http://www.aeaweb.org/aea/conference/program/retrieve.php?pdfid=406

GENDER, MEASUREMENT CHOICE AND STUDENT ACHIEVEMENT IN INTRODUCTORY ECONOMICS

David W. Brasfield, Murray State University
James P. McCoy, Murray State University
Martin I. Milkman, Murray State University

ABSTRACT

This paper uses education production functions to examine how gender influences student achievement and learning in both the Principles of Microeconomics and Principles of Macroeconomics courses at Murray State University. This study examines the influence of gender using both grades in the class and score on the Test of Understanding College Economics (TUCE) as dependent variables with the same group of students. This allows us to compare the results from regressions which are identical in every way except for different measures of student achievement and learning. Our results indicate that the choice of measurement of student achievement (standardized test or course grade) as well as the course used in the study (macro vs. micro economics) will influence conclusions concerning a gender gap in the learning of economics.

INTRODUCTION

One of the most widely studied student characteristics in terms of the learning of economics has been gender of the student. Most of the existing evidence has suggested that males generally demonstrate a higher level of economic understanding than females and that this difference first appears during adolescence (high school) or possibly even earlier. The most common explanations offered in the literature for this "gender gap" include the existence of a cultural milieu that discourages women from engaging in analytical or quantitative analysis, the relative advantage of men in spatial and numerical skill, and a classroom climate that is not conducive to female learning (Ferber, Birnbaum, & Green, 1983, p. 29).

This study explores whether conclusions about gender differences in the learning of economics are dependent upon how researchers define learning in college principles (micro and macro) courses. This is performed within the context of education production functions using two different measures of student achievement. First we use the student's grade as the measure of achievement. Given the current emphasis on student retention at most universities, the use of grades is an important consideration. However, the discrete nature of grades as an outcome measure requires the use of ordered probit analysis. Second, we use the same data set and independent variables but we use the student's score on the Third Edition of the Test of

Understanding College Economics (TUCE III) as our measure of economic achievement. This dependent variable allows for the use of the ordinary least squares estimation technique.

In each of these specifications we attempt to address the question whether the level of knowledge and understanding (a stock) at the completion of the college principles course is different for female students than for males. Also, by utilizing both pre-course TUCE scores and post-course TUCE scores, we can address the question of gender differences in the rate of learning (a flow) which occurs in the courses. Thus, by comparing the results of the two estimations, inferences can be made as to whether the answers to the two previous questions depend upon the measure of achievement utilized: grade in the class or score on a standardized test.

BRIEF REVIEW OF RELATED LITERATURE

Over thirty years ago, Siegfried (1979) completed a comprehensive survey of research involving male-female differences in student performances on economics standardized tests. Siegfried makes the useful distinction between those studies that examine differences in the stock of knowledge vs. those that study the rate of learning in a given time period (for example during a course). However, nearly all of the results reviewed by Siegfried are obtained as a by-product of research primarily concerned with other issues in economic education. Upon completing the review, Siegfried concludes that the evidence on learning and understanding economics at the elementary school level indicates few differences between females and males. However, by the high school years, gaps (in favor of males) appear to develop. Thus, he accepts as conclusive the findings of several studies that showed a gender difference in understanding of economics at the time students graduate from high school. Differences in understanding seem to persist through the college years, but there does not necessarily appear to be any widening of the gap. Specifically, Siegfried concludes that about two-thirds of the studies relating to the level of understanding which took gender into account found statistically significant higher levels of understanding for males than for females, and these tended to be studies with larger samples and more sophisticated empirical methods. On the other hand, with regard to the rate of learning in college economics, only about one-third found a statistically significant difference in favor of men.

Most of the subsequent recent research concerning gender differences has addressed whether or not the gender gap exists, and if so when does it first appear, and if it widens in college. It is probably fair to summarize this research by characterizing it as a "mixed bag." Most studies still find the existence of a gap (especially in high school or college), but several do not.

Watts (1987) provides an exception to the conclusions of Siegfried by arguing that males exhibit a higher level of economic understanding as early as grade five. Other studies involving gender differences in student performance on standardized tests of economic knowledge (not all

focus specifically on the issue of gender, but all do explicitly contain at least one independent variable to control for gender) confirm the previous findings that males tend to score higher than females (Soper & Walstad, 1988; Walstad & Soper, 1988 & 1989), especially at the college level (Ferber, Birnbaum, & Green, 1983; Gohmann & Spector, 1989; Lumsden & Scott, 1987; Watts & Lynch, 1989; Fizel with Johnson, 1986). In particular, Heath (1989) concludes that undergraduate males exhibit a higher level of economic knowledge than females, but economic learning appears to occur at the same rate in college men and women. In addition, Heath maintains that in those cases where students take economics as an elective, it is likely that only the most analytical women will choose to take the course due to cultural factors discouraging women from displaying a proclivity for quantitatively oriented courses. She concluded that self-selection does occur in choosing an economics course as an elective and that the result is a downward bias to previous estimates of gender differences in the stock of economic knowledge.

However, several studies involving student performance on standardized tests of economic knowledge have not found a gender difference (Rhine, 1989; Buckles & Freeman, 1983; Watts, 1987; Beron, 1990). In addition, when essay questions are used instead of multiple choice, Ferber, Birnbaum, and Green (1983) report that the male-female differential is reduced while Lumsden and Scott (1987) report that the nature of the gap is even reversed. Additionally, in contrast to most other studies, Lumsden and Scott (1987) report that in their study female learning rates were lower than male learning rates. Williams, Waldauer & Duggal (1992, p. 229) used both multiple choice and essay question performance to measure student achievement and "found no evidence to support the hypothesis that significant gender differences exist in college students' performance on economic exams".

Results concerning the gender gap using course grade as the measure of achievement are similarly mixed. Reid (1983) found no significant difference between males' and females' grades in an introductory college economics course as did Brasfield, McCoy, and Milkman (1992) and Brasfield, Harrison and McCoy (1992). However, Myatt and Waddell (1990) found a negative and significant relationship between being female and performance (grade) in a college introductory economics course.

One other study of note involving the issue of gender differences in economic education analyzed the role of gender in terms of enrollment in a second economics course (persistence in the study of economics). Horvath, Beaudin, and Wright (1992) concluded that female students persisted in the study of economics in lower proportions than their male classmates and those females generally required higher grades than males in order to persist in economics.

Most of the early studies described above utilize standardized economics exams (mostly multiple choice) to address issues concerning the gender gap. Ballard and Johnson (2005) in a study of microeconomics principles use grades as the independent variable, no pretest (thus a stock measurement), and find that the independent effect of gender is small and insignificant. However expected grade has a positive and significant effect on class performance and males had a significantly higher expected grade than females. Arias and Walker (2004) use the total

score on four exams given during the term. No pretest is included in their independent variables and they find that while the coefficient for females is negative, again it is not statistically significant.

In this context, this project will provide additional evidence regarding answers to the following questions with respect to possible gender differences upon the completion of college principles (micro and macro) of economics classes:

1. Is there a difference between males and females in the stock of knowledge?
2. Is there a difference between males and females in the flow of learning which takes place in the course?
3. Do the answers to these two previous questions depend upon the measure of achievement utilized (standardized test or course grade)?
4. Do the answers to these two previous questions depend upon whether macro or micro economics is the subject used in the study?

METHODOLOGY AND VARIABLE DESCRIPTIONS

The standard approach to examining the effect of gender on the learning of economics high school economic education is through the use of education production functions (Murname, 1981 and Hanushek, 1979). A total of eight education production functions are estimated, four for performance in microeconomics, and four for performance in macroeconomics. Within each set of four for each course (micro and macro), the dependent variable used in two of the education production functions is the score on the Third Edition of the Test of Understanding College Economics (TUCE III). These two specifications differ only by whether or not each student's pretest score on the TUCE (administered at the beginning of the course) is included as an independent variable. Including the pretest score as an independent variable (Specification 1) allows for inferences concerning the flow of knowledge (learning) while helping with problems of possible omitted variable bias and autocorrelation (Hanushek, 1979). Not including the pretest as an independent variable (Specification II) implies that the posttest measure is a simple measure of the stock of knowledge at the completion of the course and thus it allows for more direct comparisons with much of the previously completed research in this area. We also ran regressions under the specification that the difference between the posttest score and the pretest score on the TUCE is the dependent variable (measure of achievement). The results for both macro and micro are qualitatively identical to those reported in Tables 3 and 4 for Specification I (posttest score as dependent variable with pretest score included as an independent variable).

The technique of ordinary least squares was used to estimate Specifications I and II (standard tests indicated each specification to not be significantly affected by heteroschedasticity). The other two specifications (Specification III and IV) utilized course

grade (A, B, C, D, or F) as the dependent variable and therefore employed the ordered probit estimation technique. Again, specifications III and IV differed only by the inclusion (or not) of the pretest TUCE score. There are two versions of the TUCE: one for microeconomics and the other for macroeconomics. Each student completed the TUCE during the final exam period of their introductory economics course. The score on TUCE determined between five and ten percent of the final course grade. The TUCE pretest was administered during the first two weeks of class to measure prior knowledge. The pretest contains exactly the same questions as the final exam TUCE.

While there are some advantages to using a standardized measure of economic achievement, using grades might be "a more appropriate measure of cognitive achievement than performance on a standardized test because we are explicitly attempting to measure the ability of students to master the specific content of the course as defined by the instructors" (Bonello, Swartz, & Davisson, 1984, p. 205). For those schools emphasizing retention of students, course grade may be the relevant dependent variable, because a major factor influencing the likelihood of a student remaining in school is his or her ability to meet the expectations of instructors.

In Specifications III and IV we assume the existence of an underlying, continuous grade scale of student final grades (call it Y*) for each course, which we assume to be identical across instructors except for a shift factor which is estimated in our regression. The observed dependent variable used here, actual letter grades, thus indicates the need for ordered probit estimation where one estimates the probability of observing the qualitative letter grades.

Table 1: Variable Descriptions

POSTTEST	score on TUCE III administered during final exam
ECOGRA	grade in this course (A=4, B=3, C=2, D=1, F=0)
SEX	male/female dummy variable (1 for female)
PRETEST	score on TUCE III administered during first two weeks of class
STATE	dummy variable (1 for in-state students)
TRANS	dummy variable (1 for transfer students)
COACT	composite ACT score
MA220	completed business math sequence dummy variable (1 if completed
HSECO	dummy variable (1 if the student has had high school economics)
REGHRS	number of credit hours for which student is registered this semester
CRHRS	number of credit hours completed by student prior to this semester
COLGPA	current college GPA
REPECO	dummy variable (1 if the student is repeating this course)
OTHGRA	grade received in the other principles course (A=4, B=3, C=2, D=1, F=0)
WORK	dummy variable (1 for students working on or off campus for pay at time of survey)
ACTIVE	number of extra-curricular activities in which student is involved
HOURS	hours student plans to study in this course per week
PEER	average grade in this class section of student

A brief description of the variables used in the education production functions is presented in Table 1. The measure of student achievement is either the posttest score on the

TUCE (POSTTEST) or grade achieved in the class (ECOGRA). Several of the independent variables measure more than one influence on student achievement. For example, composite ACT scores (COACT), current college GPA (COLGPA), the student preTUCE score (PRETEST), and completed college credit hours (COLHRS) are measures of not only innate ability but also school inputs. Having successfully completed a high school economics class (HSECO), the business math sequence (MA220), or the other Principles of Economics class (OTHGRA) are more logically categorized as strictly school inputs. Regardless, we expect the sign of all of these influences to be positive. Student background and characteristics measures include the student's gender (SEX), involvement in extra-curricular campus activities (ACTIVE), and employment (WORK). These student characteristics have been included in similar previous work with mixed results. Therefore, we assign no expected sign to them. Another student characteristic variable, HOURS, measures the hours per week a student intends to study for the course as reported by the student in the initial weeks of the semester. Recognizing the limitations of this measure, we nonetheless include it with an expected positive sign. Other student characteristic variables include whether or not the student is a transfer student (TRANS), number of hours for which the student is registered (REGHRS), and whether or not the student is repeating the course (REPECO). We include these variables with no a priori expected sign. The PEER variable, defined as the average grade in the class section of the student, is included to pick up peer learning effects, but we recognize that it will pick up instructor grading biases as well. For both of these reasons we expect it to have a positive sign.

DESCRIPTION OF THE DATA

The data used in this study come from four sources. A survey was completed by students in the principles class after the last day to add the class or to drop the class without a "W" (withdrawn--no grade) appearing on the student's transcript, but before the first exam. Students completed this survey during the class period on the date the survey was distributed. The second data source is a student file maintained by the academic records office. This file was used to verify GPA and ACT statistics (for a description of issues surrounding the use of self-reported vs. independently verified data see Maxwell and Lopus (1994)).

The third data source is the final grade reports filled out by the instructors of the two principles classes. These reports contain the data for the students' course grades. Letter grades are used to evaluate student performance at Murray State by instructors and thus we were constrained to working with this discrete dependent variable. Letter grades at our institution are not qualified with pluses and minuses. Therefore, possible letter grades are A, B, C, D, and F. The fourth source is the scoring sheets for each student on both the pretest and posttest TUCE.

Table 2 contains basic information on the Principles of Microeconomics and Principles of Macroeconomics classes used in this study. The classes used in this study include principles classes offered at Murray State University for seven semesters. Though students are free to take

either the microeconomics or macroeconomics class first, most students enroll in macroeconomics before microeconomics. One likely reason for this is that the course number for macroeconomics (ECO 230) is lower than the course number for microeconomics (ECO 231).

Table 2: Descriptive Statistics for Macro Principles (based on 469 observations)

Variable	Total Mean (Std.Dev.)	Mean for Males (Std.Dev.)	Mean for Females (Std.Dev.)	Mean for Difference (t-stat.)
POSTTEST	13.307	13.531	13.070	.461
	(4.312)	(4.401)	(4.212)	(1.158)
ECOGRA	2.252	2.232	2.272	- .040
	(1.121)	(1.383)	(1.105)	(- .381)
SEX	.486	---	---	---
	(.500)			
PRETEST	9.612	9.988	9.215	.773**
	(2.886)	(3.019)	(2.688)	(2.926)
STATE	.669	.660	.680	- .020
	(.471)	(.475)	(.468)	(- .461)
TRANS	.115	.137	.092	.045
	(.320)	(.344)	(.290)	(1.524)
COACT	22.528	22.477	22.561	- .084
	(3.751)	(3.742)	(3.768)	(- .778)
MA220	.345	.303	.390	- .087*
	(.476)	(.460)	(.489)	(-1.993)
HSECO	.243	.295	.189	.106**
	(.429)	(.457)	(.392)	(2.696)
REGHRS	15.456	15.382	15.535	- .153
	(2.305)	(1.995)	(2.522)	(-.728)
CRHRS	48.699	49.299	48.066	1.233
	(24.904)	(25.475)	(24.326)	(.536)
COLGPA	2.762	2.649	2.881	- .232**
	(.645)	(.641)	(.630)	(-3.942)
REPECO	.100	.112	.087	.024
	(.301)	(.316)	(.284)	(.876)
OTHGRA	.200	.191	.210	- .020
	(.718)	(.699)	(.739)	(- .296)
WORK	.420	.357	.487	- .130**
	(.494)	(.480)	(.501)	(-2.868)
ACTIVE	1.019	1.029	1.099	.020
	(1.009)	(.946)	(1.074)	(.216)
HOURS	2.876	2.797	2.960	- .164
	(2.207)	(2.157)	(2.259)	(-.803)
PEER	2.112	1.943	1.973	- .029
	(.399)	(.458)	(.448)	(- .695)

*Indicates significance at the .05 level.

**Indicates significance at the .01 level.

Table 2: Descriptive Statistics for Micro Principles (Continued--) (based on 345 observations)				
Variable	Total Mean	Mean for Males	Mean for Females	Mean for Difference
	(Std.Dev.)	(Std.Dev.)	(Std.Dev.)	(t-stat.)
POSTTEST	13.864	14.783	13.060	1.723**
	(4.694)	(4.838)	(4.423)	(3.455)
ECOGRA	2.441	2.466	2.419	.047
	(1.071)	(1.037)	(1.103)	(.409)
SEX	.533	---	---	---
	(.500)			
PRETEST	10.423	10.484	10.370	.114
	(3.300)	(3.504)	(3.113)	(.320)
STATE	.655	.683	.630	.053
	(.476)	(.467)	(.484)	(1.028)
TRANS	.136	.174	.103	.071
	(.344)	(.380)	(.305)	(1.913)
COACT	22.235	22.590	21.924	.666
	(3.642)	(3.573)	(3.684)	(1.699)
MA220	.478	.466	.489	- .023
	(.500)	(.500)	(.501)	(- .431)
HSECO	.267	.329	.212	.117**
	(.443)	(.471)	(.410)	(2.471)
REGHRS	15.458	15.416	15.495	- .079
	(2.465)	(2.360)	(2.559)	(- .297)
CRHRS	63.089	61.217	64.726	-3.509
	(25.680)	(27.281)	(24.149)	(-1.267)
COLGPA	2.778	2.690	2.854	- .164**
	(.627)	(.686)	(.562)	(-2.441)
REPECO	.087	.099	.076	.023
	(.282)	(.300)	(.266)	(.764)
OTHGRA	1.745	1.665	1.815	- .151
	(1.364)	(1.410)	(1.322)	(-1.024)
WORK	.568	.615	.527	.088
	(.496)	(.488)	(.500)	(1.643)
ACTIVE	1.284	1.547	1.054	.492**
	(1.248)	(1.304)	(1.153)	(3.722)
HOURS	5.477	5.251	5.674	- .422
	(2.990)	(2.794)	(3.145)	(-1.31)
PEER	2.127	2.174	2.085	.088
	(.430)	(.457)	(.402)	(1.903)
*Indicates significance at the .05 level.				
**Indicates significance at the .01 level.				

Differences in the total mean values of many of the variables between macro and micro are attributed to the fact that most students take ECO 230 (macro) prior to ECO 231 (micro). The OTHGRA values in Table 2 are consistent with this typical ordering. Since the majority of students take macro first, micro students typically have completed one more semester than macro students. This extra semester of college experience for the typical micro student is evidenced by the 14.4 difference in the mean values of "COLHRS" of the two groups. It is interesting to note

that the mean values of the sex variable are higher in ECO 231 (micro) than they are in macro (230). That is, the majority (51.4%) of students in macro (typically the first class taken) are males while the majority of students (53.3%) in micro (typically the second class taken) are females. Unlike the results of Horvath, Beaudine, and Wright (1992), these data provide a less than rigorous (since not all students take macro first and some students transfer in credit for one of the courses) indication that at Murray State University females may be better "persisters" than males in principles of economics.

The difference between mean posttest and pretest scores is slightly larger in macro than it is in micro. This is probably because most students enroll in macro before micro. Mean pretest scores are higher for students in microeconomics since they have presumably learned some economics in the macroeconomics class. Also the mean value of ECOGRA (as well as PEER) is higher for the microeconomics class. This may be due to a filtering process. Those students who do not do well in the macroeconomics class may decide not to enroll in the microeconomics class.

Also included in Table 2 for both courses are means and standard deviations for each of our variables broken down by male and female samples. The final column lists differences in these means (male mean minus female mean) and the difference in means t-statistic. While not necessarily surprising, some of the significant differences are nonetheless interesting.

Males score significantly higher than females on both the pretest TUCE and the posttest TUCE in macro. Thus, males appear to begin the macro course, typically the first course taken (only 8% of the macro students in our sample had already passed micro) with a higher stock of knowledge (as measured by the TUCE) as well as end the course with a higher stock of knowledge. This is consistent with the higher proportion of males (29.5% vs. 18.9% for females) who have completed a high school economics course. However, while on average males score higher on the pretest in micro, the difference is not statistically significant. Thus, when students start micro (70% of whom have already passed macro), the gender gap as measured by the TUCE is narrowed or nonexistent. This could occur due to a relative increase in the stock of economic knowledge by females between macro and micro or a decrease in the relative stock of knowledge on the part of males. Another explanation is that the thirty percent of the students in micro who have not taken macro are responsible for this result. It is also possible that females approach the pretest in a more serious fashion than males. Regardless, at the completion of the micro course males score significantly higher than females on the TUCE for that course. Thus, the gender gap seems to reappear. Regardless, these data are consistent with regression results to be reported in the next section.

Other interesting and possibly related results become evident upon perusal of female/male differences with respect to the Composite ACT score (COACT), repeating the course (REPECO), and the grade in the other economics course (OTHGRA). While no significant difference exists between female and male ACT scores in the macro course, males in the micro course have significantly higher ACT scores. Thus some filtering may be occurring

based upon gender specific characteristics embodied in the ACT. Possibly related is the fact that the macro grades of females in micro are higher (but not significantly) than those of males. Females are also more likely to have completed the business math sequence (MA220) before enrolling in macro. (This is even more remarkable since the mean number of credit hours earned at time of enrollment in macro is greater for females than for males.) In addition, females are less likely to be repeating the micro course and possess higher college GPA's (COLGPA) than males in both courses. These results are consistent with the previously reported conclusion that "females require more concrete symbols of success (higher grades) than males in order to continue in the introductory economics sequence" (Horvath, Beaudin, & Wright, 1992, p. 107). It also seems to indicate that on average our females are "better students" with respect to course work (at least in terms of grades) than our male students. There is more to come on this later.

EMPIRICAL RESULTS

Table 3 contains the estimated education production functions for the macroeconomics class and Table 4 contains the estimates for the microeconomics class. The first column of each table presents estimates for the OLS regression using the TUCE posttest as the dependent variable and including the pretest as an independent variable (Specification I). The specification estimated in the second column (Specification II) is identical to the first except the pretest is not included as an independent variable. The third column presents estimates for the ordered probit regression with course grade as the independent variable (Specification III) with pretest included, while the specification estimated in the fourth column (Specification IV) is identical to the third except the pretest is not included as an independent variable.

Table 3: Macro Regression Results

Specification:	I: OLS Post as Dep. Variable	II: OLS Post as Dep. Variable	III: Ordered Probit Grade as Dep. Variable	IV. Ordered Probit Grade as Dep. Variable
CONSTANT	3.052	5.707**	-4.935	-4.410
	(1.759)	(3.320)	(-9.228)	(-8.688)
SEX	-0.495	-.751*	-0.239*	-0.281*
	(-1.413)	(-2.096)	(-2.078)	(-2.446)
PRETEST	.336**	---	0.059**	---
	(5.495)	---	(3.269)	---
TRANS	1.159	1.197*	-0.066	-0.057
	(2.079)	(2.080)	(- .379)	(- .322)
COACT	.339**	.401**	0.087**	0.096**
	(6.321)	(7.430)	(-5.908)	(-6.550)
MA220	-0.335	-0.44	.268*	0.248*
	(- .920)	(-1.172)	(2.275)	(2.110)
HSECO	-0.734	-0.795	-0.117	-0.125
	(-1.850)	(-1.942)	(- .925)	(- 1.002)
REGHRS	-0.132	-0.146	-0.002	-0.004
	(-1.692)	(-1.815)	(-.064)	(-.144)

Table 3: Macro Regression Results				
Specification:	I: OLS Post as Dep. Variable	II: OLS Post as Dep. Variable	III: Ordered Probit Grade as Dep. Variable	IV. Ordered Probit Grade as Dep. Variable
CRHRS	0.004	0.005	0.0003	0.001
	(0.483)	(0.659)	(0.131)	(-0.266)
COLGPA	1.019**	1.045**	1.082**	1.070**
	(-3.093)	(3.075)	(11.999)	(-12.742)
REPECO	-0.982	-1.133	-0.014	-0.042
	(-1.628)	(-1.822)	(-.079)	(-.248)
OTHGRA	0.093	0.231	0.045	0.068
	(0.379)	(0.923)	(0.529)	(-0.783)
WORK	0.01	-0.08	-0.143	-0.154
	(-0.027)	(-.222)	(-1.247)	(-1.342)
ACTIVE	-0.234	-0.296	-0.069	-0.079
	(-1.292)	(-1.590)	(-1.180)	(-1.345)
HOURS	.311**	.313**	.254**	0.252**
	(3.931)	(3.842)	(14.929)	(15.269)
PEER	-.886*	-1.149**	.665**	0.612**
	(-2.293)	(-2.903)	(5.594)	(-5.114)
MU(1)			1.29	1.279
			(10.231)	(-10.285)
MU(2)			2.694	2.663
			(18.654)	(-19.017)
MU(3)			4.201	4.148
			(23.84)	(-24.568)
R^2	0.323	0.277		
F	14.42**	12.49**		
Log-Likelihood			-500.84	-505.59
Restricted Log-L			-702.91	-702.91
Chi-Squared			404.15	394.63
Significance Level			0	0
NOTES: T-Statistics are in parenthesis.				
*indicates significance at the .05 level.				
**indicates significance at the .01 level.				

Table 4: Micro Regression Results				
Specification:	I: OLS Post as Dependent	II: OLS Post as Dependent	III: Ordered Probit Grade as Dep. Variable	IV. Ordered Probit Grade as Dep. Variable
CONSTANT	2.07	2.588	-3.524	-3.524
	-0.948	(1.152)	(-5.786)	(-5.836)
SEX	-1.998**	-1.945**	-0.178	-0.178
	(-4.512)	(-4.265)	(-1.373)	(-1.374)
PRETEST	.328**	---	.00001	---
	(4.601)	---	(.001)	---
TRANS	0.094	0.513	0.052	0.052
	(0.146)	(0.781)	(0.284)	(0.295)
COACT	.371**	.492**	.053*	0.053**
	(5.285)	(7.342)	(2.530)	(3.007)

Table 4: Micro Regression Results				
Specification:	I: OLS Post as Dependent	II: OLS Post as Dependent	III: Ordered Probit Grade as Dep. Variable	IV. Ordered Probit Grade as Dep. Variable
MA220	-0.798	-0.772	0.049	0.049
	(-1.859)	(-1.747)	(0.378)	(.378)
HSECO	-0.607	-.550	.032	-0.032
	(-1.266)	(-1.113)	(- .221)	(-.222)
REGHRS	-0.112	-0.12	.008	0.008
	(-1.260)	(-1.304)	(0.301)	(.302)
CRHRS	.019*	.024**	0.002	0.002
	(-2.246)	(2.786)	(0.791)	(0.818)
COLGPA	1.507**	1.460**	1.038**	1.038**
	(3.75)	(3.527)	(10.572)	(10.692)
REPECO	0.086	0.251	.500*	0.501*
	(0.112)	(.316)	(2.017)	(2.017)
OTHGRA	0.138	0.214	0.127**	0.127**
	(0.821)	(1.249)	(2.772)	(2.857)
WORK	-0.187	-0.378	-0.088	-0.088
	(- .436)	(- .859)	(- .665)	(- .690)
ACTIVE	-.393*	-.436*	-0.061	-0.061
	(-2.149)	(-2.317)	(-1.050)	(-1.050)
HOURS	-0.106	-0.138	-0.032	-0.032
	(-1.488)	(-1.893)	(-1.388)	(-1.395)
PEER	0.464	-0.366	0.717**	0.717**
	(- .941)	(- .721)	(4.824)	(4.865)
MU(1)			0.974	0.974
			(7.214)	(7.241)
MU(2)			2.121	2.121
			(14.067)	(14.126)
MU(3)			3.577	3.577
			(19.317)	(19.335)
R^2	0.373	0.332		
F	13.06**	11.77**		
Log-Likelihood			-405.26	405.26
Restricted Log-L			-500.02	500.02
Chi-Squared			189.51	189.51
Significance Level			0	0
NOTES: T-Statistics are in parenthesis.				
* indicates signifcance at the .05 level.				
** indicates significance at the .01 level.				

Our results lend support to the hypothesis that the existence of a gender gap in economics depend upon the specification of the model, and in our case, even the course under study (macro vs. micro). Inferences concerning the possible existence of a gender gap in favor of males appear to be linked to the various characteristics and issues embodied in the eight different regression estimates. These characteristics and issues include whether results concerning macro or micro are analyzed, what measure of achievement is utilized (TUCE vs. course grade), and whether the interest is in stock or flow of learning.

In the macro regressions (Table 3), the estimate of the gender variable (where male = 0 and female = 1) is negative but insignificant in our "value added" or flow specification (I). However, in our simple stock specification (II) the estimate of the gender variable is negative and significant. These results, in conjunction with those of Table 2, seem to indicate that females begin (using the pretest TUCE) and end the macro course (typically their first college course) with a lower stock of knowledge than males, regardless of whether we use the TUCE or course grade as our dependent variable. But given that Specification I for macro leads to no significant difference in the gender variable, when other factors are accounted for, the gap does not necessarily get larger during the macro course. This is consistent with much of the previous research. However, the results for the gender variable in the microeconomics regressions present a different picture. While similar to macro, the coefficient for sex is negative and significant when using the TUCE as our measure of achievement (for micro in both Specifications I and II), but unlike macro, the coefficient is negative but insignificant in Specifications III and IV when grades are used as the dependent variable. This result for micro using grades is consistent with the findings of Ballard and Johnson (2005). Thus, using TUCE as our measure of knowledge, females not only finish microeconomics with a lower stock of knowledge but they do so at least in part because they learn less during the course. However, they do not necessarily receive lower grades. We also estimated "gender specific" education production functions separately for males and females in our sample as a quick test for gender differences in our estimates. No remarkable qualitative differences were found. Within these "gender specific" regressions we also included a teacher-gender variable, but found no significance for this variable for males or females. It appears that the choice of dependent variable, or measurement of achievement (TUCE vs. course grade), will influence our conclusions regarding the existence of a gender gap in microeconomics (generally the last course taken here) but not in earlier macroeconomics! Table 5 summarizes the results regarding the possible existence of a gender gap in both courses.

TABLE 5

Specification:	I. TUCE with pretest	II. TUCE no pretest	III. Grade with pretest	IV. Grade no pretest
Macro Results for Student Gender	Negative but NOT Significant	Negative and Significant*	Negative and Significant*	Negative and Significant*
Micro Results for Student Gender	Negative and Significant*	Negative and Significant*	Negative but NOT Significant	Negative but NOT Significant

Tables 3 and 4 also illustrate the importance of other independent variables in predicting student achievement in both classes. Not surprisingly, students' pretest scores are significant predictors of their posttest scores in both regressions. The macro pretest score is also a significant predictor of a student's grade in macro. College grade point average (COLGPA) is a significant predictor of achievement for all specifications in both macroeconomics and microeconomics. This finding is consistent with previous pedagogical research both in and out of economics (for example see Clauretie & Johnson, 1975) and simply confirms the hypothesis

that previous achievement is a good predictor of future achievement. Similarly, coefficients for students' Composite ACT scores (COACT) are positive and significant in all specifications for both macro and micro.

The PEER variable is also significant and positive in Specification III for both courses. Simply interpreted, the better the grades of an entire section, the better an individual student is likely to do in either macroeconomics or microeconomics. This could be due to instructor grading biases (some instructors giving generally higher grades than others) or due to "true" peer learning effects. We also tried one other specification of the PEER variable: Using the mean COLGPA for each class section as our PEER variable. This alternative peer variable was not statistically significant in any of the regressions. We also tried adding instructor dummy variables. The instructor dummy variables did not change the impact of the PEER variable or the alternative specification of the PEER variable when grade was used as the dependent variable, however both the PEER variable and the alternative specification of the peer variable were insignificant when the instructor dummy variables were included. These results highlight the difficulty of separating out the peer learning effects from instructor grading biases and other section specific effects.

The math sequence variable MA220, is positive and significant in the macroeconomics regression but not the microeconomics regression. One possible explanation for this result is that principles of macroeconomics typically incorporates more algebra, especially in national income determination models, than does microeconomics. Or, possibly students learn the necessary mathematical tools in macro (typically the first course) and therefore don't need to have completed the math sequence before micro. Finally, students with poor math skills may not "persist" to micro and therefore are not present in the micro sample to make a significant difference. Oddly, the intended study hours variable is positive and significant in all specifications for macro but negative (however insignificant) in the micro regressions. Finally, one variable of interest found not to be significant in Specification III or IV of either course is number of completed college credit hours (CRHRS). This result runs counter to previous research (for example see Clauretie & Johnson, 1975), who found the same variable to be positively related and significant with course grade.

CONCLUSIONS

Our results indicate that conclusions concerning the existence of a gender gap may depend upon how the achievement variable is specified and which principles of economics course is under study. In our sample, and using the TUCE as our measure of knowledge, females appear to begin and end the college principles of economics course with a lower stock of knowledge. Results concerning a gender gap in the flow of learning however are mixed. Again, using TUCE scores as our measure of economic knowledge, the gap between males and females does not widen in our macro courses but it does widen in our micro courses. Results are also

mixed when grades are used as the measure of achievement but, with respect to macro and micro, in an opposite fashion. Specifically, females do not receive significantly lower grades in micro (when several other factors are accounted for) but they do in macro, where they begin with less knowledge perhaps due to a lack of interest or exposure to economics in high school.

Regardless of whether the perception of course grades is that they measure a stock (the more common) or a flow of knowledge, it is quite clear that the choice of measurement of achievement, or the dependent variable (TUCE score vs. course grade), as well as whether macro or micro principles of economics is the course under study, can dramatically influence conclusions concerning the gender gap (and possibly other issues as well) when estimating education production functions. If stock of knowledge is the issue, in our sample, using TUCE scores would lead to the conclusion that a gender gap exists in both macro and micro while if course grades are used, the gap only appears in macro. If flow of knowledge is the issue, using our sample's TUCE scores would lead to the conclusion that the gap exists in micro but not macro, while using grades would indicate it exists in macro but not in micro. Clearly the choice of dependent variable in estimating economic education production functions as well as the specific economics course under study (principles of macro vs. micro) is critical when addressing the role of gender in learning economics.

Unfortunately, these results lead to more questions (all of which are related) than definitive answers. Just what does constitute "knowledge" in introductory college economics? What are appropriate and accurate measures of this knowledge? Are gender differences in learning really different in macro vs. micro economics? Are measured gender differences a result of biases in the measurements themselves, or a result of different learning styles of males and females thus requiring concrete actions to improve the classroom climate for females (as suggested by Ferber 1984 and 1990)? While our estimates of simple gender specific education production functions could not provide any evidence of obvious learning style differences, we are far from convinced that this issue has been fully explored. Therefore, this area we believe to be a fruitful avenue for future research.

REFERENCES

Arias, J. & D. Walker (2004). Additional Evidence on the Relationship between Class Size and Student Performance. *Journal of Economic Education,* 35(4),311-329.

Ballard, C. & M. Johnson (2005). Gender, Expectation, and Grades in Introductory Microeconomics At A US University. *Feminist Economics*. 11(1), 95-122.

Beron, K. J. (1990). Joint determination of current classroom performance and additional economics classes: A binary/ continuous model. *Journal of Economic Education,* 21(3), 255-64.

Bonello, J., R. Swartz, & I. Davisson (1984). Freshman-sophomore learning differentials: A comment. *Journal of Economic Education*, 15(3), 205-10.

Brasfield, D., D. Harrison, & J. McCoy (1992). The impact of high school economics on the college principles of economics course. *Journal of Economic Education*, 24(2), 99-111.

Brasfield, D., J. McCoy, & M. Milkman (1992). The effect of university math on student performance in principles of economics. *Journal of Research and Development in Education*, 25(4), 240-42.

Buckles, S. G. & V. Freeman (1983). Male-female differences in the stock and flow of economic knowledge. *Review of Economics and Statistics*, 65(2), 355-58.

Clauretie, M., & E. Johnson (1975). Factors affecting student performance in principles of economics. *Journal of Economic Education*, 6(2), 132-34.

Ferber, M. A. (1984). Suggestions for improving the classroom climate for women in the introductory economics course: A review article. *Journal of Economic Education*, 15(2), 160-68.

Ferber, M. A. (1990). Gender and the study of economics. *The Principles of Economics Course: A Handbook for Instructors, Chapter 6.* New York: McGraw-Hill.

Ferber, M. A., B.G. Birnbaum, & C.A. Green (1983). Gender differences in economic knowledge: A reevaluation of the evidence. *Journal of Economic Education*, 14(2), 24-37.

Fizel, L., with D. Johnson (1986). The effect of macro/micro course sequencing on learning and attitudes in principles of economics. *Journal of Economic Education*, 17(2), 87-98.

Gohmann, S., & L.C. Spector (1989). Test scrambling and student performance. *Journal of Economic Education,* 20(3), 235-38.

Heath, J. (1989). An econometric model of the role of gender in economic education. *American Economic Review*, 79(2), 226-30.

Horvath, J., B.Q. Beaudin, & S.P. Wright (1992). Persisting in the introductory economics course: An exploration of gender differences. *Journal of Economic Education*, 23(2), 101-08.

Kmenta, J. (1986). *Elements of Econometrics*, 2d. New York: Macmillan.

Lumsden, K., & A. Scott (1987). The economic student re-examined: Male-female differences in comprehension. *Journal of Economic Education*, 18(4), 365-75.

Maxwell, N. L., & J.S. Lopus (1994). The Lake Wobegon effect in self-reported data. *American Economic Review*, 84(2), 201-05.

Murnane, R. (1981). Interpreting the evidence on school effectiveness. *Teachers College Record,* 83(1), 19-35.

Myatt, A., & C. Waddell (1990). An approach to testing the effectiveness of the teaching and learning of economics at high school. *Journal of Economic Education,* 21(3), 355-63.

Reid, R. (1983). A note on the environment as a factor affecting student performance in principles of economics. *Journal of Economic Education,* 14(4), 18-22.

Rhine, S. (1989). The effect of state mandates on student performance. *American Economic Review*, 79(2), 231-35.

Saunders, P. (1991). The third edition of the test of understanding in college economics. *American Economic Review*, 81(2), 32-37.

Siegfried, J. J. (1979). Male-female differences in economic education: A survey. *Journal of Economic Education,* 10(2), 1-11.

Soper, J. C., & W.B. Walstad (1988). What is high school economics? Posttest knowledge, attitudes, and course content. *Journal of Economic Education*, 19(1), 37-51.

Walstad, W. & J.C. Soper (1988). A report card on the economic literacy of U.S. high school students. *American Economic Review*, 78(2), 251-56.

Walstad, W. B., & J.C. Soper (1989). What is high school economics? Factors contributing to student achievements and attitudes. *Journal of Economic Education,* 20(1), 23-38.

Watts, M. (1987). Student gender and school district differences affecting the stock and flow of economic knowledge. *Review of Economics and Statistics,* 69(4), 561-66.

Watts, M., & G. Lynch (1989). The principles courses revisited. *American Economic Review*, 79(2), 236-41.

Williams, M. L., C. Waldauer, & V.G. Duggal (1992). Gender differences in economic knowledge: An extension of the analysis. *Journal of Economic Education*, 23(3), 219-31.

TICKET PRICING PER TEAM: THE CASE OF MAJOR LEAGUE BASEBALL (MLB)

Kwang Woo (Ken) Park, Minnesota State University, Mankato
Soonhwan Lee, Indiana University Purdue University, Indianapolis
Phillip Miller, Minnesota State University, Mankato

ABSTRACT

In this paper, we explore the determinants of demand for attendance at Major League Baseball (MLB) games for 23 individual MLB teams during the period 1970 to 2003. Our central focus is to explore team-specific elasticities of demand for attendance. We use Error Correction Models (ECM) to identify these elasticities. The empirical findings show that factors of demand differ between teams with respect to the factors that determine attendance and to the estimated weights. We find that demand for attendance is mostly inelastic with levels varying between teams.

INTRODUCTION

Sports teams generate revenues from three general sources: ticket sales, concession sales, and the sale of media rights. To generate maximum profits, teams must possess knowledge about the relationship between ticket prices and attendance at the team level. A host of factors influences the demand for sports, including the price of tickets, fan income, the population of the drawing area, team quality, and the age of the stadiums in which teams play. While each of these factors generally influences the demand for all teams' games in some manner, the marginal impact each has on attendance may vary between teams.

More specifically, the sensitivity of attendance to changes in the price of tickets (the elasticity of demand) and to changes in average incomes (the income elasticity) may vary from team to team. For example, some MLB teams reside in cities without NBA or NHL teams (such as Kansas City) while some reside in metropolitan areas with teams in each league. Some teams reside in cities with one or more teams from the same league, such as Chicago and New York which both have multiple teams in MLB. A large literature on the demand for sporting events exists and there have been some analyses of team-specific attendance (Simmons, 1996) and revenue (Burgers & Walters, 2003; Porter, 1992). Yet team-specific price and income sensitivity of attendance in American sports at the team level has been largely unexplored.

We attempt to fill this gap by exploring team-specific demand for 23 MLB teams. We examine time-series data that allows us to identify specific factors that affect team-specific attendance and to measure the marginal impact these factors have on attendance at the team

level. Using an error correction model (ECM), we are able to estimate the elasticities of demand and income for specific teams. To our knowledge, no other paper applies this approach in analyzing MLB data.

We organize the rest of the paper as follows: Section II presents a review of the literature; section III presents the empirical framework; section IV describes the data; section V presents the empirical results; section VI concludes.

LITERATURE REVIEW

Unit Elastic and Elastic Demand Evidence

In American professional sports, researchers generally agree that franchise owners render decisions with an eye towards maximizing profits. In addition, American franchises in the four major sports are granted exclusive territorial rights by their leagues, giving teams a measure of monopoly power in their local markets.

According to economic theory, a single-product firm will generate maximum profits when it produces an amount where the added costs of production (the marginal costs) are just equal to the added revenue from selling the product (the marginal revenue). Moreover, economists have identified a relationship between the marginal revenue and the elasticity of demand. When a firm sells more of a product, its revenue increases, all else equal. However, if the increased sales result from a lowering of the product's price, then the price change gives an offsetting effect on revenue. Do revenues increase, decrease, or remain constant? Knowledge of the elasticity of demand provides the answer to this question.

When the demand for a product is elastic, lowering a product's price causes revenues to increase (marginal revenue is positive). When the demand is unit elastic, revenues neither increase nor decrease (marginal revenue is zero). When demand is inelastic, revenues fall (marginal revenue is negative). Because a firm generates maximum profits by selling where marginal revenues equal marginal costs, a firm facing non-negative marginal costs will set its product price in the elastic or unit-elastic portion of the product demand curve.

If we assume that the marginal cost of allowing a fan into a ballpark is zero (all costs are fixed), then pricing at the unit-elastic point ensures maximum profits. Demmert (1973) examines MLB data covering the period from 1951 to 1969 and found evidence that teams set ticket prices around the unit-elastic portion of the demand curve. Noll (1974) found evidence that MLB teams priced in the inelastic portion of their demand curves on average, but he could not rule out unit-elastic pricing statistically.

Noll noted two important issues in interpreting these results. First, elasticity estimates may be understated because the price of admission is only a portion of the total costs fans pay to attend games. Second, larger parks have a greater proportion of seats with poor views than smaller parks and estimating demand equations by using simple average ticket prices (adding the

various sections' ticket prices and dividing by the number of sections) will not adequately capture the proportion of seats in different sections.

If, however, the elasticity of demand is greater than one in absolute value, demand is elastic, implying that consumers are relatively sensitive to changes in the price of the good. Alexander (2001) analyzes the demand for MLB and finds that after controlling for the price of other entertainment options, demand is elastic where prices are set.

Simmons (1996) explored the determinants of attendance by team at Premier Soccer League matches in the United Kingdom. After adjusting attendance figures for season ticket buyers and non season ticket buyers, he finds some evidence that some particular teams price in the elastic portion of their demand curves.

Inelastic Demand Evidence

The evidence that ticket prices are set in the elastic or unit elastic portion of demand is the exception, not the rule. A more consistent finding is that teams on average set ticket prices in the inelastic portion of demand. Below we summarize a set of this research. We direct readers who desire a more in-depth summary of this research to Fort (2004a).

The evidence of inelastic ticket pricing is found across sports and countries. Such pricing is found in the Scottish Football League (Jennett, 1984), in the Spanish Football League (Garcia & Rodriguez, 2002), minor league baseball (Siegfried & Eisenberg, 1980), MLB (Scully, 1989; Zimbalist, 1992), Australian Rules Football (Borland, 1987), and the NFL (Brook, 2006).

One of Noll's (1974) concerns, noted above, was that the price of tickets was only one part of the overall price of attending games. Ignoring costs of traveling to and from games, for instance, would cause overall attendance prices to be understated, thus causing estimates of the elasticity of demand to be biased downwards. Bird (1982) examined Football League attendance in England during the period 1948/49 to 1979/80 and found that after controlling for travel expenses, prices were set in the inelastic portion of demand. Forrest, Simmons, and Feehan (2002) also control for travel costs and find that previous elasticity estimates for English soccer were too low. They however could not rule out inelastic ticket pricing.

Another issue is that watching a game at a stadium is only one way for fans to "consume" the action. When franchise owners allow other avenues through which fans can follow games, they create, to some extent, a substitute for in-person attendance. Thus, whether a particular game is broadcast on television is an important determinant of the demand for tickets. Carmichael, Millington, and Simmons (1999) examine English Rugby League attendance and find evidence of inelastic ticket pricing after controlling for the telecasting of games on the British Sky Broadcasting network.

The finding of inelastic pricing has been a puzzle to economists because it suggests that teams have chosen ticket prices that are "too low" if teams seek maximum profits. If teams indeed price where demand is inelastic, raising ticket prices would generate more ticket

revenues. How can such behavior be rationalized? One possibility is that inelastic pricing promotes more people to attend games. For example, Boyd and Boyd (1998) suggest that teams would choose relatively low ticket prices to induce more fans to attend games in order to improve home field advantage. Another possibility, put forth by Fort (2004b), is that teams set lower ticket prices in exchange for public subsidies.

However, Quirk and El Hodiri (1974), Marburger (1997), and Krautmann and Berri (2007) explain that inelastic ticket pricing should be expected in certain situations because of the trade-off between gate revenue and other sources of revenue. In other words, sports teams are not single-product firms but are, instead, multi-product producers. Not only do they sell the action on the field but they also sell concessions, parking, and souvenirs. It is certainly plausible that franchise owners would happily accept lower ticket revenue in exchange for revenues from other sources.

Income elasticity evidence

Income elasticity measures the sensitivity of the quantity demanded of a product with respect to changes in consumer incomes. Estimating income elasticity not only allows for researchers to comment on whether consumers are relatively responsive to changes in income, but they also allow researchers to comment on whether a good is an inferior or normal good. While most of the focus has been on the elasticity of demand, some researchers have commented on income elasticities. Bird (1982) estimates that the demand for English soccer is inferior, suggesting that as average incomes grow, the demand for soccer will fall. However, Simmons (1996) finds that English soccer is a normal good.

The majority of attendance studies provide evidence on inelastic ticket pricing. But as noted earlier, the analyses of American sports have relied on data aggregated by league. Do all teams in a league price in the inelastic portion of their demand curves or do some price in the elastic portion of demand? An analysis of team-specific demand allows us to examine elasticities at team levels. We now move to a brief description of the theoretical framework that explains equilibrium levels of attendance and ticket prices and the empirical model.

EMPIRICAL FRAMEWORK

We postulate a long-run relationship between attendance and two generally important demand factors for each MLB team, the real price of tickets prices and real per-capita income, as follows:

$$A_{it} = \beta_0 + \beta_1 P_{it} + \beta_2 I_{it} \qquad (1)$$

where A is the logarithm of attendance, P is the logarithm of the real price of tickets, and I is the

logarithm of real per capita income in the home city. *i* and *t* denote team and year identifiers respectively.

The test for a long-run relationship described by equation (1) requires a test for the stationarity of the series. A series is said to be "stationary" if the mean, variance, and autocorrelation of the series are unchanged over time. If two (or more) series are each non-stationary but a linear combination of the series is non-stationary, then the series is said to be "cointegrated." Fortunately, equation (1) can be easily cointegrated.

It is well-accepted that habit persistence and long memory usually formulates demand functions: the current consumption of a good depends, in part, on consumers' past consumption of the good. By including lagged demand terms for attendance (Deaton & Muellbauer (1980), Borland (1987), and Simmons (1996)), the demand equation (1) can be written as an autoregressive-distributed lag model

$$\Delta A_{it} = \beta_0 + \beta_1 \Delta P_{it} + \beta_2 \Delta P_{it-1} + \beta_3 \Delta I_{it} + \beta_4 \Delta I_{it-1} + \beta_5 A_{it-1} + \beta_6 P_{it-1} + \beta_7 I_{it-1} + \varepsilon \quad (2)$$

where Δ denotes the difference operator. We can estimate equation (2) by using an error correction model (ECM). In other words, equation (2) can be reparameterized with the following ECM Structure:

$$\Delta A_{it} = \beta_0 + \beta_1 \Delta P_{it} + \beta_2 \Delta P_{it-1} + \beta_3 \Delta I_{it} + \beta_4 \Delta I_{it-1} + \beta_5 [A_{it-1} + (\beta_6 / \beta_5) P_{it-1} + (\beta_7 / \beta_5) I_{it-1}] + \varepsilon \quad (2')$$

Equation (2') is a typical ECM and the parameterization from (2) to (2') is similar to that employed by Simmons (1996). For the interpretation of short-run dynamics, we assume a stationary process in the variables given the cointegrated variables in the long-run. Thus, the differenced values in P and I are the same when the time lags are the same, and the parameters of equation (2') can be rearranged. This is the case of having a weakly stationary process in P and I. Since the cointegrating properties of the equation are not affected by reparmeterization, the equation (2') can be used for interpreting the short-run dynamics. Hence, when P and I are stationary, we can interpret $\beta_1+\beta_2$ as the short-run price elasticity, $\beta_3+\beta_4$ as the short-run income elasticity, β_6/β_5 as the long-run price elasticity and β_7/β_5 as the long-run income elasticity. In addition, the coefficient β_5 can be interpreted as the speed-of-adjustment factor for the residual terms in the stationary long-run cointegration process between the dependent variable and the explanatory variables depicted in equation (1).

We can incorporate our basic ECM into a more general ECM specification with team-specific factors that affect attendance, such as team quality. Thus, we write the general model as:

$$\Delta A_{it} = \beta_0 + \beta_1 \Delta P_{it} + \beta_2 \Delta P_{it-1} + \beta_3 \Delta I_{it} + \beta_4 \Delta I_{it-1} + \beta_5 A_{it-1} + \beta_6 P_{it-1} + \beta_7 I_{it-1} + \theta X_{it} + \varepsilon \quad (3)$$

where θ is a vector of parameters to be estimated and X_{it} is a matrix of team-specific variables. We describe these variables in the next section.

Since the estimation of equations (2) and (3) are rationalized only by establishing the existence of the long-run relationship given by equation (1), establishing cointegration between the non-stationary processes is a necessary condition for the estimation of an error correction model such as equation (3) (Engle & Granger (1991)). Without cointegration, the statistical properties in equation (3) may be spurious. Hence, we test for cointegration by using the procedure proposed by Engle and Granger (1987) and Kremers, Ericsson, and Dolado (1992) and developed by Kanioura & Turner (2003). These authors show that a cointegration test based upon a conventional F-test for the joint significance of the levels terms is advantageous because its distribution does not depend on the specific parameters of the problem being considered. We present our test results with our empirical results below.

We estimate equation (3) for each team, a process preferable to using dynamic pooled estimation. Pesaran and Smith (1995) show that if there is coefficient heterogeneity between teams, pooling the data may generate a serial correlation problem. Estimation of an autoregressive distributed lag (ADL) model with lagged dependent variables, in the presence of this serial correlation, yields inconsistent estimates. In addition, Simmons (1996) shows that fixed effects may not be sufficient to deal with these team-level idiosyncrasies. He also argues that it is unclear how to define the correct form of a pooled equation because of those idiosyncracies. Hence, estimations for separate teams allow us to capture specific team-level effects more precisely, especially those relating to team quality where the effects may vary considerably across teams. In particular, since we base our model specification upon a stationary process of error terms given long-run cointegration, we need not consider autoregressive integrated moving average (ARIMA) terms. In addition, since all the variables on the right-hand side of the equation are exogenously determined, we use simple ordinary least squares to estimate the model.

In general, stadium capacity constraints are not a problem with our data. In the 34 years of our analysis, only one of 782 total records, the San Francisco Giants in 2000, had attendance of 100% of annual capacity. A total of 10 records had attendance values at 95% of capacity or above and a total of 26 had attendance values at 90% of capacity or above. Since capacity constraints are not a problem, we do not use censored regression analysis.

We obtain the final form of model specification for each team using a 'general-to-specific' specification search, a process particularly popularized by Hendry and Mizon (Hendry, 1995; Mizon, 1995; Hendry, 1993; Hendry & Mizon, 1990). Hendry and Krolzig (2001) recommend the use of multiple search paths in the process of moving from a generalized unrestricted model (GUM) to a parsimonious specification. There are two reasons for using this process. First, it allows us to avoid deleting an important variable that we should ideally retain in the final specification along any single search path. Second, it allows us to minimize the risk that the final model is overparameterised. Therefore, we determine the final form of the general ECM equation (3) estimated for each team by parsimony, satisfactory performance against diagnostic tests incorporated with the Schwarz criterion, evidence of cointegration, and the implied long-run relation of equation (1).

DATA AND DESCRIPTION OF VARIABLES

The data comprise the 23 U.S. MLB teams that competed each year during the period from 1970 to 2003 (excluding 1989 and 1990 when no ticket price data was available). Because of the lack of sufficient data points, we exclude teams that began play after 1970 (Seattle, Florida, Colorado, Arizona, and Tampa Bay) from the analysis. We also exclude Toronto and Montreal from the analysis because of the lack of metropolitan-specific data for their respective metropolitan areas.

Table 1 summarizes the variables used in the analysis. We calculated ticket prices using weighted average ticket prices obtained from the late Doug Pappas (www.roadsidephotos.com) and from past personal correspondence with Roger Noll. Pappas and Noll each made their calculations using ticket prices by section for each team and each used the number of seats in each section as weights. In years where the Pappas and Noll data each had values for each team (1975-2005), we took the average of the values reported in each rather than choose between them. Ticket price data was only available from Pappas for 1970-1974 and 1990-2003. The Noll data only had values for 1986-1988. As noted above, neither source had ticket price data for 1989 and 1990. Therefore, we dropped records for those years from the analysis.

We obtained population and per-capita income for the U.S. metropolitan areas served by MLB clubs from the Bureau of Economic Analysis' Regional Economic Information System (REIS).

We obtained team productivity data, measured by team winning percentage, from the Lahman database (www.baseball1.com). We include each team's current and previous years' winning percentage in the regression models. The latter helps control for fan expectations about the upcoming season.

In any given year, a team in the hunt for the playoffs may attract larger attendance levels over and above that driven by a high winning percentage. To control for this effect, we include a dummy equal to one for each team that made the playoffs. Between 1969 and 1994, teams made

the playoffs by winning their division in a two-division format. In 1995, MLB installed a three-division format and teams could make the playoffs by winning their division or by winning the wild card (being the team with the best record that did not win its division). Consequently, the playoff dummy was set equal to one if a team won its division prior to 1995, to one if a team won its division or the wild card in 1995 and thereafter, and to zero otherwise. Note that the 1994 players' strike resulted in the cancelling of the playoffs that year, so no team literally won a division. However, we treated teams that led their division at the time the strike began as having won its division.

Table 1: Variable Descriptions

Variable Name	Description
C	Concession Value
A	*Log(Attendance)*
Δ	*Difference operator, eg.* $\Delta A_t = A_t - A_{t-1}$, $\Delta P_t = P_t - P_{t-1}$
P	Log(real ticket price)
I	Log(real per capita income)=log(real income/population)
W	Winning percentage
Age	Stadium age
Pf	Playoff dummy, 1=make the playoffs in the previous season
D81	Strike dummy, 1= union strike in year 1981
D94	Strike dummy, 1= union strike in year 1994
D95	Strike dummy, 1= union strike in year 1995
X	Matrix notation [W, Age, Pf, D81, D94]'
β	Regression estimators
ε	Regression residuals (error terms)

We also include dummies equal to one for each of the strike years (1981, 1994, and 1995), strikes which shortened the length of the MLB season. Schmidt and Berri (2004) show that there are no long-lasting effects of labor strikes/lockouts in U.S. professional sports: the effects of strikes are felt only during the period when the strike occurred. Consequently, we only include dummies for the strike years.

We include the age of the team's stadium in the model. This is included as a control because new stadiums present a unique draw for residents in a metropolitan area, and some will attend games to experience this new aspect of the team. Age is included quadratically in the analysis because as a stadium ages, it loses some of its newness and attendance will likely fall as a result (all else equal). However, it is plausible that an older stadium may become a historical attraction in its own right, leading to an increase in attendance once it hits a certain age. An alternative specification would be to use a dummy variable to control for new stadiums. We did not use such a variable because we wanted to estimate the gradual effect that an aging stadium has on attendance growth. In addition, stadium age will act like a time trend for teams that did not begin play in a new stadium during the sample period (such as the Boston Red Sox, the New

York Yankees, and the Chicago Cubs). Care must be taken in interpreting estimated coefficients for the age variables for these teams.

Lastly, all dollar values are expressed in constant 2003 dollars using the seasonally-adjusted consumer price index for all urban consumers obtained from the Bureau of Labor Statistics data website (stats.bls.gov).We now present the results of the empirical analysis.

EMPIRICAL RESULTS

We report the estimates obtained from a general-to-specific specification search based upon all the diagnostic tests in Table 2. Each equation does not fail the tests for first-order serial correlation, functional form misspecification, and non-normality and heteroskedasticity of residuals. We evaluated each test statistic at the 5% significance level. We tested all variables in the final equations for whether the attendance, real price and real per capita income processes are integrated of order 0 (*I(0)*) or of order 1 (*I(1)*). We verified the stationarity of all variables using the following tests: Augmented Dickey-Fuller (ADF), Phillips-Perron (PP), Kwiatkowshi, Phillips, Schmidt and Shin (KPSS), and Ng-Perron (NP). Table 2 shows the ADF test statistics for the unit root process of attendance (*A*). Each test shows that prices and incomes are *I(1)*. The attendance processes for most teams are *I(1)*, but there is uncertainty regarding the order of integrations of the following teams' attendance values: the Chicago White Sox (CHW), the Detroit Tigers (DET), the Kansas City Royals (KCR), the Los Angeles Dodgers (LAD), the Milwaukee Brewers (MIL), the New York Mets (NYM), the Philadelphia Phillies (PHI) and the Pittsburgh Pirates (PIT)). Given the sample size (32 observations), the low power of the test, and the test statistics being close to the critical values (allowing rejection of the unit root null), we can avoid the strict interpretation of the ADF test with our stationarity test of the residuals in the cointegrated estimation equation.

The test statistics from the residuals of the cointegration results are crucial to verifying the final specification for each team. The findings show that the null hypothesis of non-cointegration can be rejected for most teams at the 5% or 10% level of significance except for MIL and the Boston Red Sox (BOS). These results suggest that there is strong evidence for the existence of a long-run cointegration relationship for each team. In particular, the long-run cointegrating relationship includes the real ticket price or real per capita income except for CHW. In particular, the DW statistics are also higher than the R^2, which suggests the existence of a cointegration relationship (Sargan & Bhargava, 1983). As Sargan and Bhargava (1983) point out, DW will approach zero as the sample size increase if the residuals are non-stationary. That means that the DW statistics from the cointergrating regression can be used as an alternative cointegrating regression test.

Tables 3a and Table 3b present the results of the estimated models. The tables show that the real ticket price, real per-capita income, each team's current winning percentage, and the year dummy D1981 are the primary determinants of attendance changes common across most teams.

Most of the estimated coefficients on the price variable are negative, except for the Baltimore Orioles (BAL), BOS, and NYY. The estimates show no relationship between attendance and ticket prices for the Cincinnati Reds (CIN), the Cleveland Indians (CLE), LAD, and NYM.

The coefficients on current team winning percentage are positive and significant for all teams except the Milwaukee Brewers (MIL) and NYY. Our analysis suggests that the growth rate of attendance is not altered by changes in winning percent for MIL and NYY.

Table 2: ADF Unit root tests

Team Name	Attendance (A)
ANA	-1.75
ATL	-1.67
BAL	-1.43
BOS	-2.52
CHC	-1.76
CHW	-3.41
CIN	-2.44
CLE	-1.52
DET	-3.20
HOU	-1.89
KCR	-3.16
LAD	-3.97
MIL	-3.05
MIN	-2.06
NYM	-3.60
NYY	-1.75
OAK	-1.61
PHI	-5.20
PIT	-3.25
SDP	-2.55
SFG	0.36
STL	-2.05
TEX	-1.59
Test critical values are as follows: -3.646342 (1% level), -2.954021(5% level) and -2.615817 (10% level)	

The significant results suggest that, for example, a 1% increase in the Anaheim Angels' winning percentage would increase the growth rate of attendance by 1.98%. Winning has the largest impact on the San Francisco Giants (SFG), where a 1% increase in team win percent drives the growth rate of attendance higher by 3.46%. For the Minnesota Twins (MIN), a 1% increase in team win percent drives the growth rate of attendance higher by 3.10%. Both Atlanta (ATL) and the Chicago White Sox have estimated coefficients above 3.0, suggesting that a 1% increase in team win percent drives the growth rate of attendance by better than 3%. Winning has the smallest impact on the growth rate of attendance for BOS, where a 1% increase drives the growth rate of attendance higher by less than 1% (0.9%) on average. A potential reason for this is that the Red Sox play in one of the smallest parks in MLB. In 8 of the 34 years in our sample

period, the Red Sox had attendance levels at 90% to 98% of annual capacity. They, therefore, do not have a lot of room to grow.

Table 3a: Restricted Estimates of ECM
(Dependent Variable: ΔA_t, sample: 1970-2003)

	ANA	ATL	BAL	BOS	CHC	CHW	CIN	CLE	DET	HOU	KCR	LAD
A_{t-1}	-.84 (5.79)	-.81 (7.63)	-.81 (7.39)	-.74 (4.26)	-.87 (9.29)	-.57 (5.23)	-.56 (6.43)	-.62 (6.05)	-.65 (6.01)	-.39 (2.87)	-.94 (9.41)	-.54 (4.24)
P_{t-1}	-.74 (3.66)	-.46 (4.15)	.58 (5.91)	.42 (2.05)	-.49 (3.12)					-.90 (2.75)	-1.37 (4.62)	
I_{t-1}	1.72 (3.64)	.88 (2.65)	1.47 (3.91)	.70 (2.21)			1.09 (5.13)	1.70 (3.73)	.51 (2.48)		2.55 (4.35)	
ΔP_t					-.63 (2.29)	.92 (1.94)			.29 (1.84)		-.47 (1.77)	
ΔP_{t-1}	.98 (2.76)				.39 (1.94)						.80 (3.14)	
ΔI_t		3.52 (3.19)										
ΔI_{t-1}				-1.23 (1.26)								-1.40 (2.13)
W_t	1.98 (4.09)	3.17 (5.24)	2.04 (8.36)	.90 (.42)	1.84 (5.25)	3.01 (5.39)	1.39 (4.78)	2.85 (5.95)	1.96 (5.44)	2.32 (3.04)	2.22 (4.03)	1.48 (4.60)
W_{t-1}	1.66 (3.35)		.76 (2.24)	.72 (2.22)	1.55 (2.50)		.82 (2.96)				2.45 (5.58)	
Age				.27 (3.32)	.04 (8.45)					-.07 (2.70)	-0.03 (1.82)	.04 (3.31)
Age^2				-.002 (3.34)			-.000 (3.10)			.002 (2.68)	0.001 (1.34)	-.001 (3.11)
Pf		.31 (2.69)										
D81			-.44 (25.0)	-.63 (20.3)	-.72 (7.51)	-.40 (2.61)	-.71 (7.57)	-.54 (3.09)	-.52 (4.24)	-.57 (3.06)	-.46 (4.68)	-.34 (3.97)
D94			-.42 (18.56)	-.23 (6.99)	-.21 (2.40)	-.37 (2.16)	-.21 (2.11)		-.44 (3.61)	-.45 (2.60)	-.36 (3.35)	-.65 (4.29)
D95					-.21 (2.40)		-.25 (2.53)				-.46 (4.13)	
R^2	0.76	.89	.93	.91	.94	.71	.83	.68	.80	.62	.94	.82
DW	1.80	1.89	2.23	1.48	2.42	2.07	2.36	2.17	2.07	1.89	2.08	1.73
REST	2.17	1.64	4.52	6.23	0.09	2.53	0.05	1.66	1.35	1.12	3.39	.003
NRM	1.28	0.29	0.44	0.47	0.48	4.94	0.09	4.43	1.05	0.65	0.27	4.19
LM	1.21	.68	4.13	2.72	0.50	0.05	0.91	1.14	0.67	0.87	0.79	0.39
W	0.74	1.22	0.72	4.78	0.68	1.20	0.69	0.61	0.39	0.87	0.44	0.92
ECM	-5.38	-4.38	-6.36	-3.75	-5.96	-4.34	-6.92	-6.12	-5.49	-4.65	-4.32	-4.82

t-statistics in bracket. Δ indicates first difference. REST is Regression Specification Error Test proposed by Ramsey (1969). NRM is the Jarque-Bera statistic for testing normality. LM is a Lagrange multiplier (LM) test for autoregressive conditional heteroskedasticity (ARCH) in the residuals (Engle 1982). W is a test for heteroskedasticity in the residuals from a least squares regression (White, 1980). ECM is the test statistics for Engle-Granger cointegration test. The critical values are -5.75 at 1%, -4.53 at 5% and -.3.99 at 10%.

Table 3b: Restricted Estimates of ECM
(Dependent Variable: ΔA_t, sample: 1970-2003)

	MIL	MIN	NYM	NYY	OAK	PHI	PIT	SDP	SFG	STL	TEX
A_{t-1}	-.97 (8.52)	-.96 (11.71)	-.55 (6.31)	-.59 (5.18)	-.46 (4.81)	-.76 (11.02)	-.80 (8.00)	-.84 (6.66)	-.65 (5.91)	-.68 (6.21)	-.91 (8.53)
P_{t-1}	-1.11 (4.13)	-.44 (1.81)		.43 (1.78)	-1.01 (-4.16)	-.32 (3.10)	-.51 (1.75)	-1.11 (2.80)	-.40 (2.75)	-.40 (3.21)	-.19 (1.97)
I_{t-1}	2.43 (7.38)	1.25 (5.27)	.47 (2.75)		-1.74 (1.72)		1.36 (3.87)		1.28 (3.76)	1.35 (5.32)	1.75 (7.49)
ΔP_t											
ΔP_{t-1}											
ΔI_t					-2.83 (2.19)	-1.87 (1.93)					
ΔI_{t-1}											
W_t		3.10 (5.65)	1.95 (4.61)		2.68 (5.48)	1.80 (6.32)	1.93 (3.40)	1.66 (3.07)	3.46 (6.15)	2.40 (5.29)	2.43 (7.44)
W_{t-1}	2.57 (4.65)	1.68 (3.38)									.63 (1.84)
Age				.15 (3.75)	0.05 (2.45)		-.05 (2.73)				-.08 (6.04)
Age^2	-.001 (5.36)	-.001 (7.34)		-.001 (3.61)			.001 (2.58)	.001 (3.75)			.002 (5.67)
Pf				.15 (2.47)							-.24 (3.01)
D81	-.62 (4.08)		-.52 (3.23)	-.51 (15.0)		-.46 (4.71)	-.75 (6.96)	-.97 (8.82)	-.63 (3.60)	-.70 (5.49)	-.53 (5.88)
D94			-.54 (3.46)	-.55 (13.4)				-.55 (7.53)		-.29 (2.34)	-.51 (4.38)
D95				-.38 (5.85)				-.56 (6.24)			
R^2	.84	.82	.71	.73	.70	.87	.78	.87	.81	.81	.88
DW	2.45	1.88	2.28	2.29	2.78	2.15	1.90	1.71	1.64	1.71	2.47
REST	4.25	1.72	3.65	0.15	0.30	1.43	0.41	3.70	2.28	1.26	0.04
NRM	0.24	0.17	0.01	1.57	0.27	0.21	0.19	0.05	6.56	1.11	1.44
LM	1.59	0.62	1.06	3.18	2.27	0.75	0.35	1.03	0.60	1.02	2.44
W	1.22	0.40	3.90	0.53	2.68	1.27	0.83	0.51	0.53	1.16	1.34
ECM	-3.64	-4.77	-6.41	-5.75	-7.53	-4.59	-5.75	-4.62	-4.47	-4.65	-6.78

t-statistics in bracket. Δ indicates first difference. REST is Regression Specification Error Test proposed by Ramsey (1969). NRM is the Jarque-Bera statistic for testing normality. LM is a Lagrange multiplier (LM) test for autoregressive conditional heteroskedasticity (ARCH) in the residuals (Engle 1982). W is a test for heteroskedasticity in the residuals from a least squares regression (White, 1980). ECM is the test statistic for Engle-Granger cointegration test. The critical values are -5.75 at 1%, -4.53 at 5% and -.3.99 at 10%.

The coefficients on lagged winning percentage are positive and significant for 9 of the 23 teams. The results suggest that lagged winning percentage has the largest influence on the

attendance growth rates of ANA, CHC, KCR, MIL, and MIN. For example, if the lagged win percent of MIN increases by 1%, its attendance growth rate the following year rises by 2.57%. If the lagged win percent of KCR increases by 1%, its attendance growth rate rises by 2.45%. Of the teams with significant coefficients, lagged win percent has the smallest effect on attendance growth rates for Baltimore (BAL), the Red Sox, the Reds (CIN), and Texas (TEX). If lagged win percent grows by 1%, each of these sees the growth rate of attendance increase by less than 1% on average.

The coefficients on the playoff dummies were significant for only a handful of teams (the Atlanta Braves (ATL), NYY, and the TEX. The results suggest that making the playoffs caused attendance growth to rise by 0.31% the following year in ATL and by 0.15% for NYY. The sign on the estimated coefficient on the playoff dummy for TEX is unexpected since it suggests that attendance changes were smaller when the Rangers made it to the playoffs after controlling for other factors.

The age of the stadium is important in determining the change in attendance in just over half of the models. Since we entered stadium age quadratically, not only do the models tell us about the effect of the age of the stadium on attendance changes but also on the rate of change of the impact of stadium age. If a team played in the same stadium during the entire sample period, then the age variable coincides with a time trend. Teams that played in the same stadium from 1970-2003 are Anaheim (ANA), Boston (BOS), the Chicago Cubs (CHC), the Los Angeles Dodgers (LAD), the New York Mets and Yankees (NYM and NYY respectively), Oakland (OAK), Philadelphia (PHI), San Diego (SDP), St. Louis (STL). Care must be taken in interpreting the results of the age variable for these teams. A negative linear term along with a positive quadratic term suggests that as a stadium ages, attendance changes fall but the rate of decrease diminishes as the stadium ages. This finding is consistent with the honeymoon effect documented by Leadley & Zygmont, 2005; Clapp & Hakes, 2005). Indeed, for HOU, KCR, PIT, and TEX, the teams for which a honeymoon effect is present, attendance growth rates peak when their stadiums were 17.5, 15, 25, and 20 years of age respectively. Also note that each of these teams had new stadiums built during the sample period, so the age of the stadium does not correspond with a time trend.

Positive linear terms along with negative quadratic terms tell us that as a stadium ages, attendance changes rise but at a decreasing rate. This latter nature is exhibited for some teams that play in classic stadiums: BOS, LAD, and NYY. It is possible that in terms of increasing the growth rate of attendance, these classic stadiums have some historical value to fans. However, the coefficients suggest that attendance growth rates peaked at 67.5, 20 and 75 years of age respectively for these three teams. Since each of these teams played in the same stadium throughout the sample period, the age of the stadium corresponds with a time trend, so these results must be used with care.

Indeed, this interpretation of historical value is consistent with models that show negative linear terms with positive quadratic terms. Together these estimates tell us that as new stadiums

grow older, attendance changes fall but that the fall subsides up until some point where historical interest begins to take over. Of course if all stadiums were classic, then it is quite possible that the historical value might be subject to diminishing marginal utility. We can say something similar about newer stadiums.

The parameter estimates for the strike-year dummy variables for 1981 and 1994 have the expected signs and are both statistically significant for most teams. In particular, the 1981 strike shows significant effects on 19 teams' attendance growth, all except for ANA, ATL, MIN and OAK. Additionally, the magnitude exceeds 0.5 in absolute value for 14 out of those 19 teams. In other words, attendance growth rates were lower by over 0.5% for these teams. In particular, the Padres (SDP) were most affected in 1981 as their growth rate in attendance fell by nearly 1%. In 1994, the Yankees (NYY), the Padres, the Mets (NYM), the Rangers (TEX), and the Dodgers (LAD) all saw their attendance growth rates shrink by at least 0.5%. This is expected because both years were shortened because of the strikes. Only a handful of teams (CHC, CIN, KCR, NYY, and the San Diego Padres (SDP)) had lower attendance changes in 1995 as a result of the 1994 strike. The dummy for 1995 is significant for only 5 of the 23 teams studied. Since both 1994 and 1995 were shortened because of the players' strike, the change in attendance between those years is explained by the other factors for most teams.

Table 4 presents the long-run price elasticities calculated from the results. Consistent with others' findings (summarized by Fort (2004)), more teams exhibit inelastic rather than elastic demand. The results suggest that the size of the long-run price elasticities varies considerably across teams. Some teams such as KCR, MIL, OAK and SDP have long-run elastic demand indicating that lowering ticket prices would lead to increased ticket revenue. However, most teams' long-run price elasticities are significantly less than 1 in absolute value, suggesting that those teams price in the inelastic portion of their demand curves. In particular, two teams, TEX and PHI, have long-run price elasticities lower than 0.5. Recall that Quirk and El Hodiri (1974), Marburger (1997), and Krautmann and Berri (2007) argued that inelastic pricing should not be surprising when teams receive revenues from sources other than ticket sales. While our data do not allow us to examine ancillary revenue sources by team, it is possible that KCR, MIL, OAK, and SDP rely more-heavily on the sales of tickets to generate revenue compared to teams such as PHI and TEX that have estimated price elasticities well below 1.

The estimated signs of the long-run price elasticities are negative in every case except for BAL, BOS and NYY. BOS and NYY are classic rivals with a large legion of loyal fans. Moreover, both teams played in classic ballparks during the entire sample period. It is possible that our specifications do not adequately control for these effects. Baltimore, on the other hand, was the first franchise to move into a new "retro" stadium. Although we control for the newness of stadiums in our models, we may not adequately control for the uniqueness of Baltimore's stadium in particular.

The long-run income elasticity is also presented for 15 teams in Table 4. In all but one case, the estimate is positive and greater than one indicating that baseball attendance is income

elastic and a normal good for fans in those particular teams' cities. Some teams, including BOS, DET, and NYM are shown to have demand that is income inelastic, meaning that attendance is relatively insensitive to changes in per-capita income. The estimated average long-run income elasticity is 1.88 across all teams for which we give values, indicating that for every 1% increase in per capita income the attendance has gone up by 1.88 percent. Baseball in Oakland, according to the results, is an inferior good.

Table 4: Long-run Elasticities of Attendance

Team	Price	Income
ANA	-0.88	2.05
ATL	-0.57	1.09
BAL	0.72	1.81
BOS	0.57	0.95
CHC	-0.56	
CHW		
CIN		1.95
CLE		2.74
DET		0.78
HOU	-2.31	
KCR	-1.46	2.71
LAD		
MIL	-1.14	2.51
MIN	-0.46	1.30
NYM		0.85
NYY	0.73	
OAK	-2.20	-3.78
PHI	-0.42	
PIT	-0.64	1.70
SDP	-1.32	
SFG	-0.62	1.97
STL	-0.59	1.99
TEX	-0.21	1.92
Average	-0.67	1.41

For calculating the long-run elasticities, refer to the equation (2')

CONCLUSION

Our purpose in this paper is to investigate the team-specific demand for attendance for the 23 MLB clubs that competed each year between 1970 and 2003 and to estimate the long-run elasticity of demand for each team. We find that factors such as the price of tickets, the level of per-capita income in a team's host city, the team's current winning percentage, and dummy

variables that control for strike periods were significant factors in explaining changes in attendance. Other factors that affect the demand for some team's games are the age of the stadium, the previous performance of the team, and whether the team made the playoffs. We also find that the weights given to these factors vary from team to team. We also find evidence for both elastic and inelastic pricing of tickets at the team level although most teams price in the inelastic portion of their demand curves. This finding is consistent with the notion that sports teams in general are not single-product firms with market power but are, instead, producers of multiple products.

The policy implications of our work are as follows: if a team prices tickets in the inelastic portion of its demand curve, then to increase its overall revenue, it can either increase ticket prices or it can generate offsetting revenue from the sales of concessions, souvenirs, parking, or other ancillary products. If a team prices its tickets in the elastic portion of its demand curve, then as long as the costs of serving fans at the margin is close to zero, profits can be raised by decreasing ticket prices.

REFERENCES

Alexander, D. L. (2001). Major League Baseball: Monopoly pricing and profit maximizing behavior. *Journal of Sports Economics*, **2**, 341-355

Bird, P. (1982). The demand for league football. *Applied Economics,* **14**, 637-649.

Borland, J. (1987). The demand for Australian rules football. *The Economic Record,* **63**, 220-230.

Boyd, D., & Boyd, L. A. (1998). The home field advantage: Implications for the pricing of tickets of professional team sporting events. *Journal of Economics and Finance,* **22** 169-179.

Brook, S. L. (2006). Evaluating inelastic ticket pricing models. *International Journal of Sport Finance*, **1**, 140-150.

Burger, J. D., & Walters, S. J. K. (2003). Market size, pay and performance: A general model and application to Major League Baseball. *Journal of Sports Economics*, **4**, 108-125

Cairns, J., Jennett, N., & Slone, P. (1986). The economics of professional team sport, a survey of theory and evidence. *Journal of Economic Studies,* **13**, 3-80.

Carmichael, F., Millington, J., & Simmons, R. (1999). Elasticity of demand for rugby league attendance and the impact of BskyB. *Applied Economic Letters*, **6**, 797-800.

Clapp, C. M., & Hakes, J. K. (2005). "How long a honeymoon? The effect of new stadiums on attendance in Major League Baseball." *Journal of Sports Economics,* **6**, 237-263

Deaton, A., & Muellbauer, J. (1980). *Economics and consumer behavior*. Cambridge University Press, Cambridge.

Demmert, H. (1973). *The economics of professional team sports*. Lexington, MA: D.C. Heath & Co.

Engle, R. (1982) Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica*, **50**, 987-1008.

Engle, R., & Granger, C. (1987). Cointegration and error correction: Representation, Estimation, and Testing, *Econometrica*, **55**, 251-276.

Engle, R., & Granger, C. (1991). *Long-run economic relationships: Readings in cointegration*. Oxford University Press, Oxford.

Forrest, D., Simmons, R., & Feehan, P. (2002). A spatial cross-sectional anaysis of the elasticityof demand for soccer. *Scottish Journal of Political Economy*, **49**, 336-355

Fort, R. (2004a). Inelastic sports pricing. *Managerial and Decision Economics*. **25**, 87-94.

Fort, R (2004b), Subsidies as incentive mechanisms in sports, *Managerial and Decision Economics*, **25**, 95-102

Garcia, J., & Rodríguez, P. (2002). The determinants of football match attendance revisited: Empirical evidence from the Spanish Football League. *Journal of Sports Economics,* **3**(1), 18-38.

Hendry, D. (1993). *Econometrics: Alchemy or science?* Oxford: Black Well Publishers.

Hendry, D. (1995). *Dynamic econometrics*. Oxford University Press.

Hendry, D., & Krolzig, H. (2001). *Automatic econometric model selection using PcGets*. London: Timberlake Consultants Ltd.

Hendry, D., & Mizon, G. (1990). *Procrustean econometrics: Or stretching and squeezing data, In: C. Granger (ed.), Modeling Economic Series: Readings in Econometrics Methodology*. Oxford University Press.

Jennett, N. (1984). Attendance, uncertainty of outcome and policy in Scottish league football. *Scottish Journal of Political Economy,* **31**, 176-198.

Kanioura, A., & Turner, P. (2003). *The error correction model as a test for cointegration*. University of Sheffield Working Paper.

Krautmann, A. and Berri, D. (2007). Can we find it in concessions? Understanding price elasticity in professional sports. *Journal of Sports Economics*, **8**, 183-191

Kremers, J., Ericsson, N., & Dolado, J. (1992). The power of cointegration test. *Oxford Bulletin of Economics and Statistics*, **54**, 325-348.

Leadley, J. C., & Zygmont, Z. X. (2005), "When is the honeymoon over? National Basketball Association attendance 1971 – 2000," *Journal of Sports Economics,* **6**, 203-221.

Marburger, D. (1997). Optimal ticket pricing for performance goods. *Managerial and Decision Economics*, **18**, 375 – 381.

Mizon, G. (1995). *Progressive modeling of macroeconomic time series: The LSE methodology*. EUI Working Paper, ECO No 95/10, European University Institute.

Noll, R. (1974). *Attendance and price setting. In R. Noll (Ed) Government and the sports business.* Washington, D.C.: Brooklings Institution.

Pesaran, H., & Smith, R. (1995). Estimating long-run relationships from dynamic heterogeneous Panels. *Journal of Econometrics,* **68**, 79-130.

Quirk, J., & El Hodiri, M. (1974). *The economic theory of a professional sports league. In R. Noll. (Ed.) Government and the sports business.* Washington, D.C.: Brooklings Institution.

Ramsey, J. (1969). Tests for specification errors in classical linear least squares regression analysis. *Journal of the Royal Statistical Society*, Series B, **31**, 350-371.

Sargan, J., & Bhargava, A. (1983). Testing residuals form least squares estimators for being generated by the Gaussian Random Walk. *Econometrica*, **51**(1), 153-174.

Schmidt, M., & Berri, D. (2004). "The impact of labor strikes on consumer demand: An application to professional sports." *American Economic Review*, **94**(1), 344-357.

Scully, G. (1989). *The business of major league baseball*. Chicago, IL: University of Chicago Press.

Siegfried, J., & Eisenberg, J. (1980). The demand for minor league baseball. *Atlantic Economic Journal,* **8**, 59-69.

Simmons, R. (1996). The demand for English league football, a club-level analysis. *Applied Economics,* **28**, 139-155.

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

Zimbalist, A. (1992). *Baseball and billions: A probing look inside the big business of our national pastime. New York*: Basic Books

ACCESSIBILITY OR ACCOUNTABILITY? THE RHETORIC AND REALITY OF NO CHILD LEFT BEHIND

David R. Aske, University of Northern Colorado
Laura S. Connolly, University of Northern Colorado
Rhonda R. Corman, University of Northern Colorado

ABSTRACT

Can school choice and school accountability truly leave no child behind? Politically and socially popular beliefs in the power of the free market have led to a movement towards accountability and quality assurance that relies on the powers of competition. The No Child Left Behind Act (NCLB) promotes the idea that competition between schools will increase the efficiency and effectiveness of the education system. At the same time, the rhetoric of NCLB maintains the progressive message of the "Common School" era. Specifically, the forms of school choice and school accountability are at odds with the concept of universal provision of education outlined in the verbiage and title of NCLB. This article employs a model of an education production function to explore the dichotomy existing between the rhetorical intent and practical implications of the NCLB. The analysis centers on the classic efficiency/equity trade-off to show that NCLB is leading to an educational environment attempting to reach two conflicting and incompatible goals. Federal, state, and local policymakers must confront this incompatibility in order to design a policy that reflects the values most preferred by society.

The cultural faith in the power of free market competition has led to a movement towards accountability and quality assurance in the provisioning of education. The No Child Left Behind Act of 2001 (NCLB) promotes the idea that competition between schools will increase the efficiency and effectiveness of the education system. At the same time, the rhetoric of NCLB maintains the progressive message of the "Common School" era. The competitive aspects of school choice and school accountability are at odds with the concept of universal provisioning of education outlined in the verbiage and title of NCLB.

This article employs economic theory to explore the dichotomy existing between the rhetorical intent and practical implications of the NCLB. The analysis centers on the classic efficiency/equity trade-off and shows that NCLB creates an educational environment deadlocked in a battle with itself over how to reach two conflicting and incompatible goals.

AMERICAN CULTURAL VALUES OF EDUCATION

The idea that all children in the United States have the right to a publicly supported education regardless of race, social class or religious beliefs is an American value. Not only access to a public education, but the expectations of a common educational experience, is part of the American culture. This common school idea is based on the view that education should be an equitable, assimilative, and inclusive institution designed to prepare students to be future productive citizens (Meyer, 2006).

The development of the common school ideal has its roots in the nineteenth century rural, one room school house (Pulliam and Van Patten, 1999). These schools were funded by local property taxes, free to all (white) children, and governed by the local communities with little state regulation. Schools and the education students received were seen as products representing the community. Today public schools are still financed through local property taxes (although states, and to a lesser extent the federal government, do provide funding), are still open to all (all) students and governed by local school boards. And, not unlike the nineteenth century school, today's public school is seen as a representation of the community. Arguably, the public school is more locally entrenched and community based than any other economic, social, or political institution.

While historians date the end of the common school era in the United States at the end of the nineteenth century, the common school ideal remains. The rhetoric of twentieth century education policy, through Supreme Court decisions and federal legislation, reiterates the importance of attempting to achieve social equity through public education. The history of American education is rife with changes; changes in the role of the Federal government, in curriculum, in funding, in assessment, just to name a few. However, the common theme, at least in the rhetoric, is that public education in America provides all children with a "level playing field".

The 1954 Brown *v*. Board of Education decision reiterates the importance of the common school ideal as an American value. The court found the separate but equal clause of the Plessy *v*. Ferguson case in violation of the 14th Amendment. In their decision, the court made clear the importance of equality in public education, equality for all. Writing the Court's opinion, Chief Justice Warren stated:

> Today, education is perhaps the most important function of state and local governments. Compulsory school attendance laws and the great expenditures for education both demonstrate our recognition of the importance of education to our democratic society. It is required in the performance of our most basic public responsibilities, even service in the armed forces. It is the very foundation of good citizenship. Today it is a principal instrument in awakening the child to

> cultural values, in preparing him for later professional training, and in helping him to adjust normally to his environment. In these days, it is doubtful that any child may reasonably be expected to succeed in life if he is denied the opportunity of an education. Such an opportunity, where the state has undertaken to provide it, is a right which must be made available to all on equal terms (Brown *v*. Board of Education, 347 U.S. 483 , 1954).

Another Supreme Court decision, which further illustrates the common school ideal as a fundamental aspect of the American character comes from the 1963 Abbington School District *v*. Schempp case. Justice Brennan, writing a concurrence to the court's opinion, stated:

> It is implicit in the history and character of American public education that the public schools serve a uniquely public function: the training of American citizens in an atmosphere free of parochial, divisive, or separatist influences of any sort - an atmosphere in which children may assimilate a heritage common to all American groups and religions (Abbington School District v. Schempp, 372 U.S. 241, 1963).

The most expansive Federal legislation regarding public education was the Elementary and Secondary Education Act (ESEA) of 1965. This legislation was a major component of President Johnson's "War on Poverty". Congress has reauthorized ESEA eight times since 1965, the No Child Left Behind Act of 2001 was one such reauthorization of ESEA. The most significant provision of ESEA is Title I. Title I provides funds to school districts with high concentrations of economically disadvantaged children. Title I's statement of purpose reads:

> The purpose of this title is to ensure that all children have a fair, equal, and significant opportunity to obtain a high-quality education and reach, at a minimum, proficiency on challenging State academic achievement standards and state academic assessments (No Child Left Behind Act of 2001, Title I).

Again, the common school ideal remains in the rhetoric of the most important piece of Federal legislation regarding public education. Title I reflects the Federal government's commitment to equality for all, in this case, through funding.

While the common school ideal remains an important American cultural value, another important American cultural value, freedom, has become a part of the discussion of public education. (Bartlett, Frederick, Gulbrandsen, and Murillo, 2002) Regarding public education, freedom is realized through school choice. The ability of parents to decide what school their children will attend is freedom extended to education. The momentum of choice over the past few decades has been so pervasive that most parents now take choice for granted. While

historically freedom has always been an important aspect of the American culture, freedom as reflected in public school choice is a relatively new phenomenon. There is an extensive history of private school education in the United States and parents have historically "chosen" schools by locating their household in a certain city or neighborhood. However, intra-district open enrollment, inter-district open enrollment, charter schools, and home schooling as public school choice options only began entering state policy at the end of the 1980s (Colvin, 2004).

NCLB legislated school choice at the Federal level. Regarding the array of choices for parents under NCLB, the U.S. Department of Education website states:

> The *No Child Left Behind Act* provides new education options for many families. This federal law allows parents to choose other public schools or take advantage of free tutoring if their child attends a school that needs improvement. Also, parents can choose another public school if the school their child attends is unsafe. The law also supports the growth of more independent charter schools, funds some services for children in private schools, and provides certain protections for homeschooling parents. Finally, it requires that states and local school districts provide information to help parents make informed educational choices for their child (No Child Left Behind Act of 2001).

Title V of NCLB requires local school districts to provide parents with information so that they may make well-informed decisions regarding what school to send their children. Much of the information parents receive is also legislated by NCLB in Title VI. NCLB requires states to administer standardized tests to assess student achievement. The results of these tests are typically aggregated at the school level. While different states have different tests and different formats to present information to the public regarding student achievement, the information parents receive is used as a measure of how well students at a specific school are performing.

The intent of NCLB regarding school choice and accountability is clear: provide information to parents regarding the performance of the school their child is attending (using a variety of metrics), as well as performance information regarding other schools their child *could* be attending. NCLB has, therefore, created a basis and framework for competition between schools. Parents look primarily at the test scores of students at various schools to see which school is the performing the best (based on higher test scores), and for many parents the scores will influence their choice of schools. Assessment and accountability drive choice. Choice is freedom; a very central American value.

The most far-reaching impacts of NCLB on the public education landscape are the Federal mandates to state governments regarding: the assessment of student achievement through standardized testing (assessment); the provision of information regarding student/school performance (accountability); and introduction of legislation that provides parents with options regarding which school their children may attend (choice). Can the ideal of the common school

be maintained within this context of assessment, accountability, and choice? This is the fundamental question studied in this paper.

ECONOMIC NATURE OF EDUCATION

Economists have long discussed the dichotomous nature of the public and private sectors and therefore the realms in which each should engage. Further, there has been much discussion regarding where the line separating the two realms should be drawn. It is the spirit of those discussions that leads the following debate regarding the nature of public education as a good.

In pure economic terms, private goods are those that the private or for-profit sector of the economy willingly produces since that production offers opportunity to earn profit. These goods exhibit two defining traits within their nature: rivalry and excludability. Rivalry arises when the consumption of the good by one patron decreases the remaining supply of that good for other patrons. Excludability arises when a patron is prevented from receiving benefits from the good if they have not paid for the privilege. Within the context of these defining traits, education is a private good: it is rival (one additional student within the classroom will decrease the amount of personal attention received by other students in that same classroom) and excludable (if you have not paid the school's tuition, you may not attend). In the context of this simple definition, the provision of education should lie in the realm of the private sector. However, it is often argued that education generates positive externalities, thereby justifying public involvement in correcting the market failure. Much of the historical intent of the "common school" and subsequent judicial and legislative actions regarding the importance of education to the proper functioning of society reflect the externalities resulting from education.

In addition to the externality issue that leads to public sector intervention, there also arise issues related to the "proper" distribution of educational opportunities. As discussed in the previous section, education, to Americans, has not been a commodity that is available only to those with the means and the will to pursue it. Public education is funded primarily through the taxation of real property; and property ownership is highly correlated with higher levels of wealth. Providing educational opportunities to children of non-property owner families basically involves a redistribution of income.

In his seminal work, *Equality and Efficiency: The Big Tradeoff*, economist Arthur Okun describes inefficiencies associated with redistributive activities in terms of a "leaky bucket", wherein he states that when transferring income from wealthy individuals to poorer individuals, a portion of the income is lost in the process. The reasons for the leaky bucket as identified by Okun, include administrative costs associated with the redistribution and behavioral changes induced by the redistribution. These behavioral changes impact work effort; savings and investment decisions; and attitudes and motivations toward acquiring human capital (Okun, 1975). The value that society places on a more equitable distribution is illustrated by their willingness to forgo some level of efficiency to achieve it.

THEORY

In this section, we formally model the trade-off between equity and efficiency in education. Stigliz (1974; 2000) argues that the gain from education can be measured in terms of productivity. Admittedly, there are many benefits from education beyond its effect on productivity, but it is instructive to begin by restricting the model to a one-dimensional outcome measure. Other outcomes, and how they relate to the efficiency/equity trade-off, are discussed in later sections.

Stiglitz (1974) defines the education production function for student *i* as $\theta_i m(x)$, where θ_i measures the difference in ability across individuals and $m(x)$ is a function mapping a level of education spending, x, into a given level of productivity. It is assumed that there are positive but diminishing returns to education spending throughout, i.e., $m' > 0$ and $m'' < 0$.

Stiglitz's formulation allows only for differences in the marginal effect of education on productivity. We add a "shift" parameter, α_i, to allow for differences in the initial level of productivity for an individual who has no formal education. Specifically, our production function is $\alpha_i + \theta_i m(x)$, where $m(0) = 0$. Both the initial advantage and the marginal advantage may be due to either innate or environmental factors (or a combination of both).

In later work, Stiglitz (2000) discusses the trade-off inherent in the allocation of a fixed level of education spending. He defines "compensatory education" as that level of spending for which productivity is equalized across the groups. Compensatory education requires spending a larger proportion of the fixed education dollars on the "less able" individuals, in order to compensate them for their less advantageous starting point.

Compensatory education relies on an outcome-based assessment of student achievement, specifically measured in productivity, and does not address the equality of pedagogical quality (or educational inputs) used to achieve that level of productivity. The distinction between compensatory education and equalization of education expenditure is important. As Stiglitz observes, either of these measures may be viewed as the equitable policy. The difference between these two views is based on whether one believes government should attempt to equalize inputs (expenditure) or outputs (achievement). While we do not take a direct stance on which is a better measure, we do note that NCLB's Title VI focuses on accountability, which in practice has employed outcome based measurements of student achievement to assess school quality. It is our contention therefore that the goal of Title VI is to ensure some minimum degree of compensatory student achievement. Stiglitz does argue that under certain conditions (which we make more explicit below), there exists a trade-off between efficiency (defined as maximum output when summing over all students) and either measure of equity.

For simplicity, we restrict our analysis to two types of individuals, who we call "A" and "B." We assume that type "A" has an advantage over type "B" due to natural, familial, and/or environmental differences. This advantage may manifest itself through differences the initial

level of productivity, $\alpha_A > \alpha_B$, or through differences in the marginal productivities, $\theta_A > \theta_B$, or both. We assume, without loss of generality, that $\alpha_A = \alpha \geq 0$ and $\alpha_B \equiv 0$ throughout the analysis, meaning that type "A" may be able to achieve a positive level of productivity in the absence any educational spending and group "B" always requires some minimum amount of spending to achieve a positive level of productivity.

EQUITY AND EFFICIENCY CONDITIONS

Compensatory Education: The first condition of interest is the compensatory level. In this case, spending must ensure equal productivity across groups. Letting A^c and B^c be the compensatory spending levels for groups A and B, respectively, the required condition for compensatory education is:

$$\alpha + \theta_A m\left(A^c\right) = \theta_B m\left(B^c\right) \tag{1}$$

Equal Expenditure: The equal expenditure condition simply requires that spending is the same for each group. Letting A^e and B^e represent these levels, the condition is:

$$A^e = B^e \tag{2}$$

Efficiency (Pareto Optimality): Efficiency requires that funds are spent so as to maximize total productivity summed across both types. Let G be total government expenditures on public education. Since the model is formulated in terms of education spending, one additional dollar spent on group A means one less dollar spent on group B. In other words, the slope of the education "budget line" is -1. Thus we have:

$$\max_{A,B}\left(\alpha + \theta_A m(A) + \theta_B m(B)\right) \; s.t.\; A + B \leq G$$

The Lagrangian function is:

$$L = \alpha + \theta_A m(A) + \theta_B m(B) + \lambda(G - A - B)$$

The first order conditions (assuming all allocated money is spent) are:

$$\theta_A m'(A) - \lambda = 0$$
$$\theta_B m'(B) - \lambda = 0$$
$$G - A - B = 0$$

Thus, the key condition needed for efficiency is:

$$\frac{\theta_A m'(A^o)}{\theta_B m'(B^o)} = 1 \tag{3}$$

where A^o and B^o are the Pareto Optimal levels of spending on groups A and B, respectively.

COMPARISON OF SPENDING LEVELS

We now compare these spending levels under different assumptions in order to show the conditions under which the efficiency and equity goals of NCLB are incompatible. In each of the figures below, three representative isoquants are shown. Each isoquant represents all combinations of spending on groups A and B that results in equal social productivity.

Each figure also shows the "compensatory path," the "efficiency path," and the "equal expenditure path". The compensatory path connects the combinations of spending on each group needed to achieve equal total productivity, while the efficiency path connects the efficient spending combinations. The equal expenditure path is the 45-degree line. Note that total productivity summed over all students rises as we move along any of the three paths.

We consider three cases: a baseline case, in which both groups are identical; the case in which group A has a higher initial productivity ($\alpha > 0$); and the case in which group A has a higher marginal productivity ($\theta_A > \theta_B$). The situation in which group A has an advantage in both the initial and marginal productivities is simply an aggregate of the previous two.

Baseline Case: The baseline case is established by assuming the two groups are identical: $\alpha = 0$ and $\theta_A = \theta_B = \theta$. In this case, it is easy to see that spending levels coincide for all three conditions. Since the functions are the same, using equal spending for each group [condition (2)] equalizes total productivity [condition (1)] and also ensures the ratio of the marginal products equals 1 [condition (3)]. This case, shown in Figure 1, is useful for understanding how the conditions diverge once the assumption of identical groups is relaxed.

Figure 1: Both groups are identical

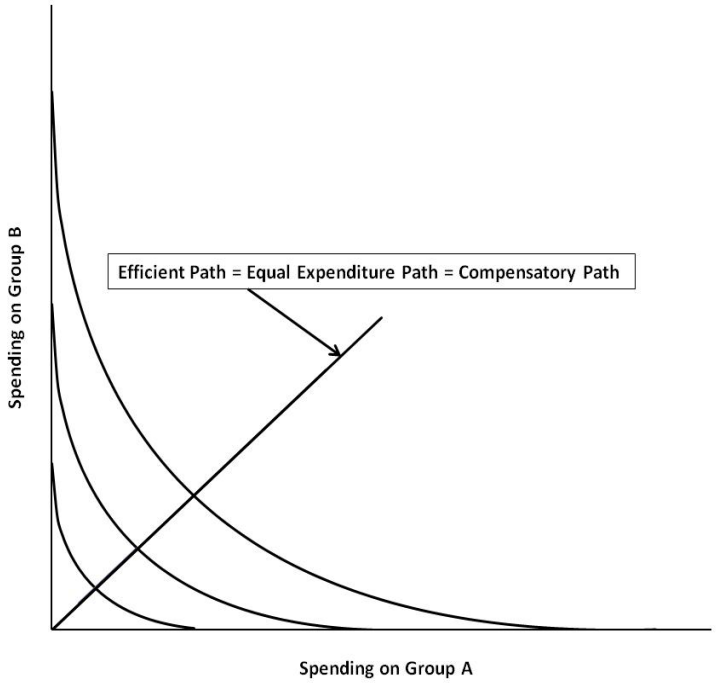
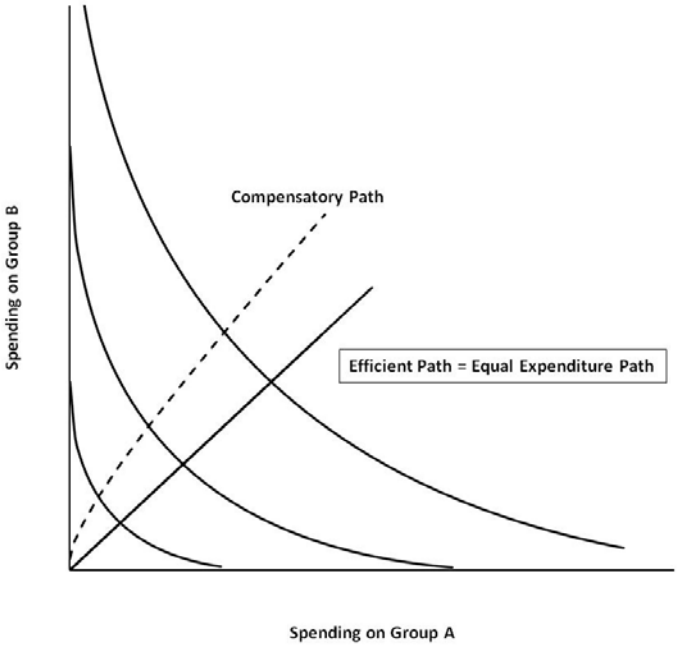


Figure 2: Group A has higher initial productivity

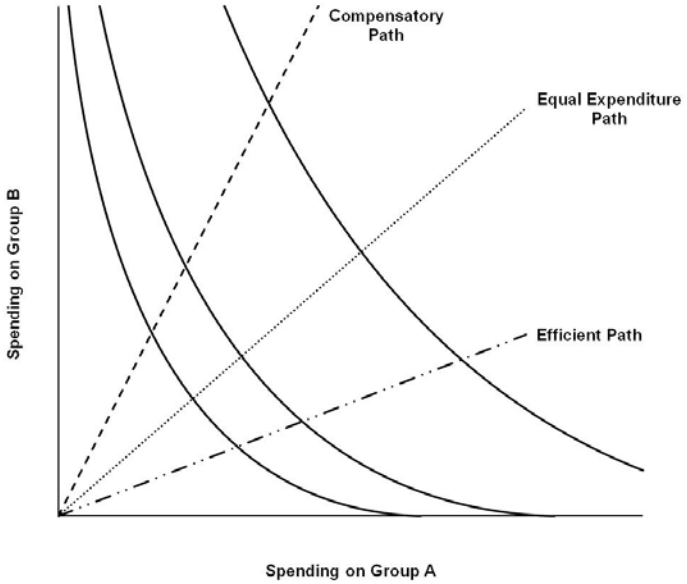


Case I: The first case of interest assumes the two groups have the same marginal productivity, but group A has a higher initial productivity level: $\alpha > 0$ and $\theta_A = \theta_B$. In this case, the isoquants are shifted outward compared to the baseline case (because the same level of total spending results in higher total productivity summed over all students) but the slopes of the isoquants remain the same (because the ratios of marginal productivities are still equal). However, compensatory education will require more spending on group B in order for them to "catch up" to group A. Indeed, this is where the term "compensatory" comes from. Spending thereafter must remain higher for group B to keep the total productivity levels of the two groups equal. Thus, the compensatory path lies above the equal expenditure path. Since the marginal productivities of the two groups remain constant, the efficient path and the equal expenditure path still coincide. This is shown in Figure 2 above.

Case II: Another case of interest arises from the assumption that the two groups have the same initial productivity level, but the marginal return to education is higher for group A, that is: $\alpha = 0$ and $\theta_A > \theta_B$.

The first thing to note is that $\theta_A > \theta_B$ implies that the isoquants will be steeper than in the baseline case because the marginal rate of substitution is $\frac{\theta_A m(x)}{\theta_B m(x)} = \frac{\theta_A}{\theta_B} > 1$. This is shown in Figure 3.

Figure 3: Group A has higher marginal productivity



Also shown in Figure 3 are the three paths of interest. As can be seen from condition (1), compensatory education will again require spending more on group B, the disadvantaged group. In this case, the compensatory path intersects the origin because we assume that only the marginal productivities differ: the initial productivities are equal in this case. The compensatory path is also steeper than the equal expenditure path for two reasons: (1) group B initially has lower marginal productivity than group A, and (2) there are diminishing returns to education, so it takes more and more additional spending on group B relative to group A to achieve equal productivity levels for each group.

Alternatively, efficiency requires that more be spent on group A in this situation. This is clear from condition (3), which requires that $m'(A^o) < m'(B^o)$ when $\theta_A > \theta_B$. Recalling our assumption that there are diminishing returns to education, $m'' < 0$, the optimal level of spending on group A must be higher than the equal spending level while the spending on group B must be lower, that is, $A^o > B^o$.

This clearly illustrates the trade-off between equity and efficiency. In both Case I and Case II, compensatory education requires a higher level of spending on group B relative to group A than would be efficient. If the two cases were to be combined, meaning that group A had an advantage in both parameters, this divergence is amplified. Thus, if compensatory education is used as the metric of equity, it is impossible to achieve equity and efficiency simultaneously. Note that if there are marginal differences, then even equality of spending is inefficient. Thus, the desired objectives of NCLB are inherently mutually exclusive.

RESULTS AND DISCUSSION

The analysis presented has shown that the objectives set forth in NCLB are ultimately incompatible; that the ensuing tradeoff between equality and efficiency makes creation of a policy that meets all of our cultural goals virtually impossible. Embedded within the American culture are the ideals of individual freedom and equal opportunity for all. To fully implement these ideals and truly realize the full extent of these convictions we, as a society, must closely examine and define the associated implicit parameters and prioritize these in the formation of policy which can strike a balance or compromise.

Much of the controversy within the current education debates have to do with the level of the contents of Okun's bucket and more specifically, how to measure the flow rates that affect the bucket's level. That is, how much is gained by investing in less-advantaged children compared to the loss resulting from redistribution? Explicit costs associated with education are easy to track but the implicit external benefits that accrue to society are much harder to assess. Blank (2002, p. 464) states, "when all children are in mandatory public schooling it is hard to measure the effects relative to a world with no public schooling, to determine the long-term returns on public school dollars."

CONCLUSIONS

The No Child Left Behind Act attempts to achieve both equity and efficiency but we have shown it cannot do both. The desire of the policymakers who developed NCLB to capture both American ideals is admirable but it is important for stakeholders (e.g., policymakers, education officials, and parents) to confront this dilemma honestly, so that an appropriate balance can be struck. This balance can only be found by conscientiously acknowledging the trade-off and by understanding both the short-run and long-run consequences that result.

REFERENCES

Abbington School District *v*. Schempp, 372 U.S. 241 (1963)

Bartlett, L., Frederick, M., Gulbrandsen, T., & Murillo, E. (2002). The Marketization of Education: Public Schools for Private Ends. *Anthropology & Education Quarterly, 33*(1), 1-25.

Blank, R. M. (2002). Can Equity and Efficiency Complement Each Other? *Labour Economics, 9*, 451-468. Retrieved from http://www.elsevier.com/locate/econbase

Brown *v*. Board of Education, 347 U.S. 483 (1954)

Colvin, R. L. (2004). Public School Choice: An Overview. In F. M. Hess & C. E. Finn Jr. (Eds.), *Leaving No Child Behind? Options for Kids in Failing Schools* (pp. 11-36). New York, NY: Palgrave Macmillan.

Meyer, H.-D. (2006). The Rise and Decline of the Common School as an Institution: Taking "Myth and Ceremony" Seriously. In H.-D. Meyer & B. Rowan (Eds.), *The New Institutionalism in Education* (pp. 51-66). Albany, NY: State University of New York Press.

No Child Left Behind (NCLB) Act of 2001, Pub, L, No, 107-110, 115 Stat, 1425 (2002, January 8) Retrieved October 20, 2009 from: http://www.ed.gov/policy/elsec/leg/esea02/index.html.

Okun, A. M. (1975). *Equality and Efficiency: The Big Tradeoff*. Washington, D.C.: The Brookings Institution.

Pulliam, J. D., & Van Patten, J. (1999). *History of Education in America* (D. A. Stollenwerk, Ed., 7th ed.). Upper Saddle River, NJ: Prentice-Hall, Inc.

Stiglitz, J. (1974). The Demand for Education in Public and Private School Systems. *Journal of Public Economics*, *3*(4), 349-385. Retrieved from EconLit database.

Stiglitz, J. E. (2000). Education. In *Economics of the Public Sector* (3rd ed., pp. 420-448). New York, NY: W. W. Norton & Company, Inc.

EMPIRICAL INVESTIGATION AND MODELING OF THE RELATIONSHIP BETWEEN GAS PRICE AND CRUDE OIL AND ELECTRICITY PRICES

Morsheda Hassan, Wiley College
Raja Nassar, Louisiana Tech University

ABSTRACT

Crude oil and natural gas are the main sources for energy in the US and around the word. Natural gas is a relatively clean source of energy compared to oil and could be cheaper to the consumer than oil especially if there is no coupling in price between oil and gas. Therefore, it is of interest to determine the long term relationship between oil and gas prices and to develop a model for predicting gas price. In this study, we used the Johansen integration test and showed that the logarithm of crude oil prices and the logarithm of natural gas prices are co-integrated in the sense that they are in a long term equilibrium relationship in which case the two series stay together and do not diverge over time. Any divergence is usually short term and eventually the two series come back together. Also, the logarithm of electricity price was found to be co-integrated with the logarithm of natural gas price. The logarithm of the GDP was also found to be co-integrated with that of natural gas price. Using data from 1973 to 2009, a time series model was developed that related the logarithm of natural gas price to that of crude oil and electricity prices. The model is useful for predicting, in the short run, gas price from knowledge of oil and electricity prices

INTRODUCTION

The United States as well as the rest of the world are heavily dependent on oil for their energy requirement. In the US, there has been an on- going call and a recent movement on the part of the government for developing alternative, renewable, and clean energy sources. A salient argument in favor of this move is to reduce our dependence on foreign oil for our own security and economic well being and to reduce environmental pollution. A sharp increase in crude oil prices over a short time period, caused by reduced oil production due to political instability in some of the oil rich countries, could cause inflation in the US and jeopardize its security.

Developing alternative clean energy sources such as wind, bio-fuel, solar and hydro power is a long term development that should be pursued. However, an immediate source of energy that is available in abundance domestically is gas. Natural gas is believed to be an important, if not the most important, energy source for the future. Natural gas is a relatively clean source of energy compared to oil and could be cheaper to the consumer than oil especially

if there is no coupling in price between oil and gas. Hence, there has been some interest in the literature in looking at the relationship of crude oil and natural gas prices. An understanding of this relationship would help in market forecasts and utilization of both commodities.

In this paper, we use time series data of crude oil and natural gas prices on a yearly basis in the US to investigate the long term relationship between them as well as between natural gas price and electricity price. In addition, we develop an empirical model relating natural gas price to crude oil price and price of electricity

RELEVANT LITERATURE

There are a number of studies in the literature investigating the relationship between natural gas price and other energy prices such as oil and electricity. Asche et al (2006) studied the long term relationship between oil, gas and electricity prices. The authors' interest was in determining if prices of any two series are co-integrated in the sense that the two series stay together and do not diverge over time. Any divergence is usually short term and eventually the two series come back together. If two series are co-integrated, then they are deemed coupled.

There was evidence from the study pointing to a co-integrated relationship between energy prices in the United Kingdom (UK) for the period, 1995-1998. Also, Pangiotidis and Rutledge (2006) found evidence for co-integration between oil and gas prices in the United Kingdom in the period 1996 to 2003. Bachmeier and Driffin (2006) presented evidence for weak co-integration between natural gas, crude oil, and coal process in the US. They argued that weak integration between oil and gas prices may be caused by the fact that only in limited areas of energy utilization, such as residential and commercial heating, there is competition between oil and gas. Barcella (1999) found a co-integrated relationship between oil and gas prices in the US which was attributed to long-run economic factors. In addition, there was a high correlation of 0.916 between yearly prices of oil and natural gas.

Serletis and Herbert (1999) , using a short time period of only one year, found co-integration between fuel oil and natural gas prices, but not between natural gas and electricity prices. They developed univariate and bivariate models to determine the relations between energy prices.

De Vany and Walls (1999) reported on co-integration between electricity prices in eleven regional markets in the US. Likewise, Hendry and Juselius (2000, 2001) found co-integration between weekly gasoline prices in different regional markets.

Brown and Yuecel (2007) using the Johansen method for co-integration and an error-correction model showed that natural gas prices and crude oil weekly prices in the US were co-integrated over the period 1994-2006. This was attributed to the fact that natural gas and crude oil are used as substitutes in energy consumption. Also, it was reported that oil price, as the independent variable, had an influence on gas price, as the dependent variable. On the other hand, gas price as an independent variable has no significant effect on crude oil price. In other words, causality was only from oil to natural gas.

DATA

Data for the United States on gas price (residential, cents per 1000 cubic feet) , oil price (Crude oil domestic, cents per barrel), electricity price (Retail price, hundredth cent per kilowatt-hour), average yearly temperature, coal consumption (residential sector, 1000 short ton), and GDP for the year 1973 to 2009 were obtained from different sources, Energy Information Administration (www.eia.doe.gov), National Climate Data Center (www.ncdc.noaa.gov/), and Economic Time Series Page (www.economagic.com).

Plots of the natural gas price, crude oil price, and electricity price for the years 1973-2009 are presented in the Appendix. These were the three variables represented in the time series model.

METHODOLOGY

The SAS software was used in the data analysis. The johansen co-integration analysis was performed in order to determine if co-integration exists between gas price and oil price and between gas price and electricity price for the US data. Also, time series transfer function analysis was used to build an empirical model relating gas price to oil price and electricity price.

Co-integration

Two time series variables are co-integrated if they possess a long-run equilibrium relationship, in which case the two series stay together and do not diverge over time. Any divergence is usually short term and eventually the two series come back together. Furthermore, two series that are co-integrated may or may not be correlated in the short-run.

Table 1 presents the co-integration analysis results for oil, gas and electricity using the Johansen co-integration test, Johansen (1988).

It is seen that for oil and gas there is no indication that the two are co-integrated (both rank = 0 or 1 are not for null and alternative hypotheses cannot be rejected since the trace value is less than the critical value for rejection at the 5% level). However log(oil) and Log(gas) seem to be integrated since the trace value is larger than the critical value when the rank is 0, but less than when the rank is 1. This says that there is long-term linear relationship between the logarithms of oil and gas. There was evidence, using the Box-Cox transformation technique, that a log transformation may be required in order to stabilize the variance. This is so in spite of the fact that the series as seen in Fig. 1 do not show strong fluctuation over time. The fact that oil and gas are only co-integrated when using a variance stabilizing transformation may indicate that the co-integration is not strong enough to manifest itself without the transformation. Weak co-integration, as pointed out by Bachmeier and Griffin (2006), may be caused by the fact that only

in limited areas of energy utilization, such as residential and commercial heating, there is competition between oil and gas.

Results from Table 1 show that log(gas) and log(electricity) are co-integrated. Also gas and electricity, without the log transform, are co-integrated. The co-integration between gas and electricity may be due to the fact that natural gas is used for residential heating and for generating electricity for residential use.

The GDP time series was found also to be integrated with gas. This was true under a log transform. Although the two series were integrated they were not correlated or functionally related as evident from the fact that the time series transfer function analysis did not show any significant relationship between log GDP and log gas. This can occur since co-integration does not imply correlation.

Table 1: Johansen co-integration Rank test for oil, gas and electricity prices.

Variables	H_0 : rank = r	H_a: rank > r	Trace	Critical Value
Oil, gas	0	0	9.27	12.21
	1	1	2.04	4.14
Log (oil), log(gas)	0	0	23.21	12.21
	1	1	2.23	4.14
electricity, gas	0	0	17.94	12.21
	1	1	1.37	4.14
Log(elec), Log(gas)	0	0	21.49	12.21
	1	1	2.57	4.14

Transfer function time series modeling

In addition to the long-run co-integration question, it is of interest to determine a short run functional relationship between gas price and other independent variables such as oil, electricity, temperature, GDP and coal. GDP, coal, and temperature have been cited among the variables that could have an effect on gas prices (US Energy information release, 2010). A strong economic growth can cause an increase in demand for natural gas and therefore a higher price. Coal and natural gas markets can interact since they can be used interchangeably for energy and electricity generation. Cold temperatures can influence residential and commercial demands for natural gas and influence prices.

For time series model, the best modeling approach is to use the transfer function analysis approach (Wei, 2006) where an output series (in this case natural gas price) is related to one or more input series. This approach is especially relevant when there is no feed back between the output and input series as determined by the cross-correlation function. If the cross-correlation between two stationary series is significant for only zero and positive lags, then there is no feed back between the output and input series (Wei, 2006). This was the case for the series considered in this study.

A transfer function model between two series, y and x, can be expressed as

$$Y_t = v(B)x_t + e_t \qquad (1)$$

where e_t is noise series that is independent of x_t.

Here, $v(B) = \sum v_j B^j$,where B is the backshift operator, $Bx = x_{t-1}$.

The function v(B) is determined from the cross correlation between x and y.

The steps involved in the identification of the transfer function model are (Wei, 2006):

1. Prewhitening of the input series.

$$\phi (B)_x = \Theta_x(B)\, \alpha_t$$

or

$$\alpha_t = \phi (B)_x / \Theta_x(B) \qquad (2)$$

2. Compute the filtered output series

$$\beta_t = (\phi (B)_x / \Theta_x(B))\, y_t \qquad (3)$$

3. Calculate the sample cross correlation between α_t and β_t

($\rho_{\alpha\beta}$ (k)) so as to determine v_k, where

$$V_k = \rho_{\alpha\beta} (k)\, (\sigma_\beta/\sigma_\alpha) \qquad (4)$$

4. Identify v(B)

Match the pattern of v(k) with the known theoretical patterns 0f v(B) in order to identify v(B).

Once v(B) is identified, express e_t in Eq. (1) as

$$e_t = y_t - v(B)\, x_t \qquad (5)$$

and identify the appropriate model fro Eq. (5) to determine the final model in Eq.(1).

In building the transfer function model, we used the logarithm of oil price differenced twice, the log of natural gas price differenced twice, the log of electricity price differenced twice,

the log of GDP differenced twice, coal consumption differenced twice and temperature differenced once.

This was necessary in order to render each of these series stationary as determined from the dampening pattern of the autocorrelation (Wei, 2006) and the Dickey-Fuller test for unit root (Dickey and Fuller, 1979).

It was found that GDP, Coal consumption and temperature were not related to natural gas price. The cross-correlation between natural gas price and each of the three series above was not significant and the coefficient estimates in the transfer function model were also not significant. As a result, the final model we present expresses the log of natural gas price as a function of the log of oil price and the log of electricity price.

The model satisfied the diagnostic checking, namely

1. There was no cross correlation between the noise series and the independent or input series (log-oil and log-electricity), which indicates that the error was independent of the input series (Chi-squared tests gave p = 0.296 and 0.297 for crude oil price and electricity price, respectively.
2. Both the autocorrelation function, ACF, and the partial autocorrelation function, PACF, for noise in the model showed no pattern. Also a formal chi-squared test up to lag 6 confirmed that the noise was white noise (p=0.116).

As a result, the transfer function model below was an adequate model relating the log of natural gas price to that of oil and electricity.

$$\begin{aligned} lgp\,(1,1)_t &= -0.39829\, lgp(1,1)_{t-1} + 1.02977\, lep(1,1)_t - 0.48577\, lep(1,1)_{t-1} - 0.35683\, lep(1,1)_{t-2} \\ &\quad + 0.14329\, lcop(1,1)_{t-1} + 0.05707\, lcop(1,1)_{t-2} + {}^{at} \end{aligned} \qquad (6)$$

where lgp (1,1) is the logarithm of natural gas price differenced twice, lep(1,1) the logarithm of electricity price differenced twice, and lcop(1,1) the logarithm of crude oil price differenced twice. Here, a_t is white noise.

FORECASTING

In using the model in (6) for forecasting lgp_t, one may replace $lgp(1,1)_t$ in Eq. (6) by

$(lgp_t - 2lgp_{t-1} + lgp_{t-2})$, (where lgp is the logarithm of the observed natural gas price) and

likewise for lep and lcop.

This gives

$$lgp_t - 2lgp_{t-1} + lgp_{t-2} = -0.39829\,(lgp_{t-1} - 2lgp_{t-2} + lgp_{t-3}) + 1.02977\,(lep_t - 2\,lep_{t-1} + lep_{t-2}) - 0.48577\,(lep_{t-1} - 2lep_{t-2} + lep_{t-3}) - 0.35683(lep_{t-2} - 2lep_{t-3} + lep_{t-4}) + 0.14329\,(lcop_{t-1} - 2lcop_{t-2} + lcop_{t-3}) + 0.05707\,(lcop_{t-2} - 2lcop_{t-3} + lcop_{t-4}) + a_t \quad (7)$$

In order to obtain the forecast for lgp_t at time t, one needs to predict the value for lep_t.

This can be obtained from the following time series model fitted to the observed values of the time series lep(1,1):

$$Lep(1,1)_t = -0.34123\, lep(1,1)_{t-4}$$

Or

$$(lep_t - 2\,lep_{t-1} + lep_{t-2}) = -0.34123\,(lep_{t-4} - 2lep_{t-5} + lep_{t-6}) \quad (8)$$

In Eqs. (7) and (8) t-k is observed data.

The models in Eqs. (7) and (8) were calculated based on the data with the last observation deleted (year 2009). Then, the model was used to predict the last observation for 2009. The predicted log of gas price was changed to gas price by taking its anti-log.

Table 1: Observed and predicted values for natural gas prices when the model was fitted to the data with the last 3 years, 2 years, or 1 year deleted.

	Data used:	Last 3 years deleted		Last 2 years deleted		Last year deleted	
Year	Observed	Predicted	Difference	Predicted	Difference	Predicted	Difference
2007	1308	1440	132				
2008	1389	1494	105	1294	95		
2009	1356	1535	179	1261	95	1438	82

As expected, it is seen that the model prediction (when the last 3 years were deleted from the data) was better for 2007 and 2008 then for 2009. However, all three predictions were not significantly different from their observed values at the 5% significance level.

When the model was fitted to the data, excluding years 2008 and 2009, the model predictions for 2008 and 2009 improved. The difference between observed and expected was reduced to 95. Also, the model fitted to the data with 2009 deleted, gave the best prediction for 2009 where the difference was reduced to 82. These results are consistent with what one expects in that prediction becomes less accurate for the distant future. In practice, one should use the model to predict the next year for a more reliable forecast.

CONCLUSION

Results of this study are useful in that they show a long-term relationship between the logarithm of oil and gas prices as well as the logarithm of gas and electricity prices. This is in

agreement with what has been postulated in the literature. The fact the gas and oil prices when not transformed were not related or co-integrated may indicate a weak relationship. On the other hand, the co-integration between gas and electricity prices seem to be stronger in the sense that it held under untransformed as well as transformed data. A time series model was developed, based on data between 1973 and 2009, which related the output series (gas price) to the input series, crude oil price and electricity price. Eq. (7) shows that first, second, and third lags were involved in the functional relationship between gas and oil and electricity. The model was able to predict gas price adequately from knowledge of oil and electricity prices.

The most reliable approach for forecasting is to update the model when data become available and use it to predict the value for the next year in the future.

REFERENCES

Asche, F. , P. Osmundsen and M. Sandsmark (2006). The UK market for natural gas, oil and electricity: are the prices decoupled? The Energy Journal 27, 27-40.

Bachmeir. L. J. , and J. M. Griffin (2006). Testing for market integration: Crude oil, coal and natural gas. The Energy Journal 27, 55-71.

Barcella, M. L. (1999). The pricing of gas. Oxford Energy Forum. Oxford Institute for Energy Studies, Oxford. May.

Brown, S. P. A. and M. K. Yuecel (2007). What drives natural gas prices? Research Department, Federal Reserve Bank of Dallas, P.O. Box 655906, Dallas, TX 75265.

De Vany, A. and W. D. Walls (1999). Cointegration analysis of spot electricity prices: insights on transmission efficiency in the Western US. Energy Economics 21, 435-448.

Dickey, D. A. and W. A. Fuller (1979). Distribution of the estimates for autoregressive time series with a unit root. .J. Amer. Statist. Assoc. 74, 427-431.

Hendry, D. and K. Juselius (2000). Explaining cointegration analysis: Part I. Energy Journal 21, 1-42.

Hendry, D. and K. Juselius (2001). Explaining cointegration analysis: Part II. Energy Journal 22, 75-120.

Johansen, S. (1988). Statistical analysis of cointegration vectors.Journal of Economic Dynamics and Control 12, 231-254.

Panagiotidis, T. and E. Rutledge (2007). Oil and gas markets in the UK: Evidence from cointegration approach. Energy Economics 29, 329-347.

Serletis, A. amd J. Herbert (1999). The message in North American energy prices. Energy Economics 21, 471-483.

U.S. Energy Information Administration (2010). Factors affecting natural gas prices. Department of Energy. February issue.

Wei, W.W.S. (2006). Time series analysis: Univariate and Multivariate methods. Pearson Addison Wesley, New York.

APPENDIX

Figure 1: Plot of electricity price over years

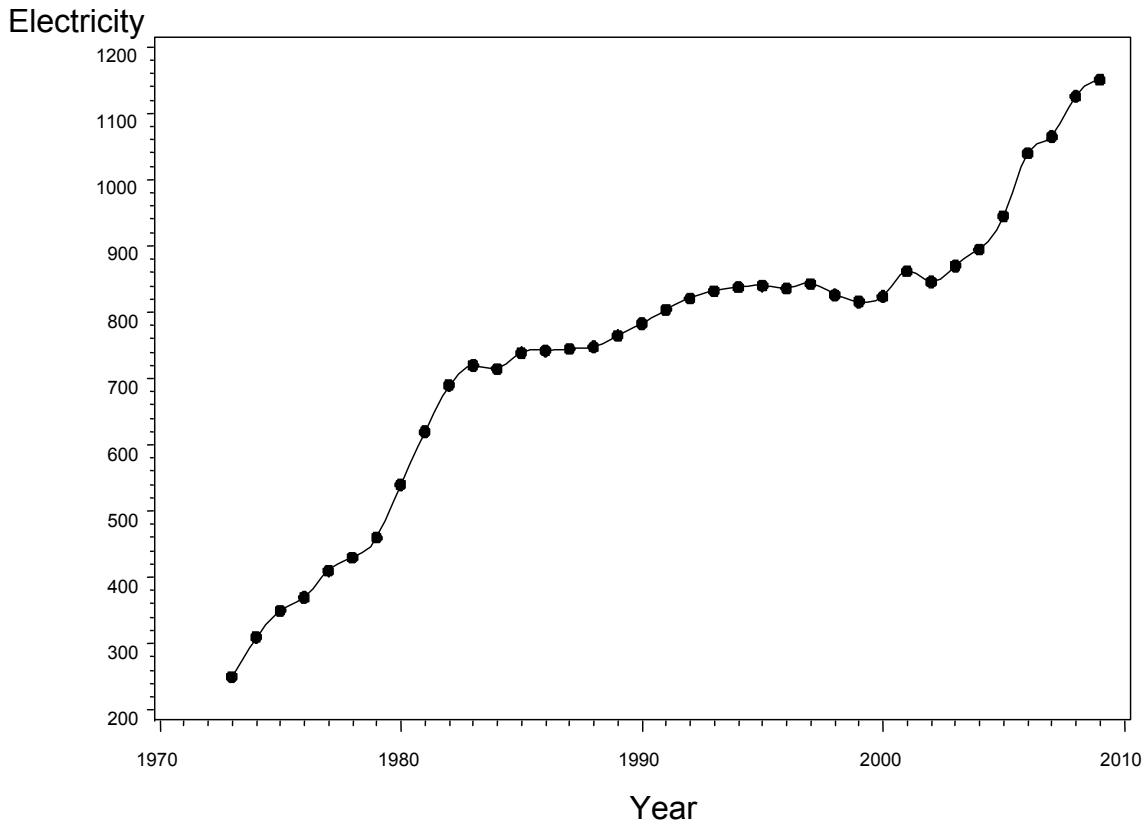


Figure 2: Plot of crude oil price over years.

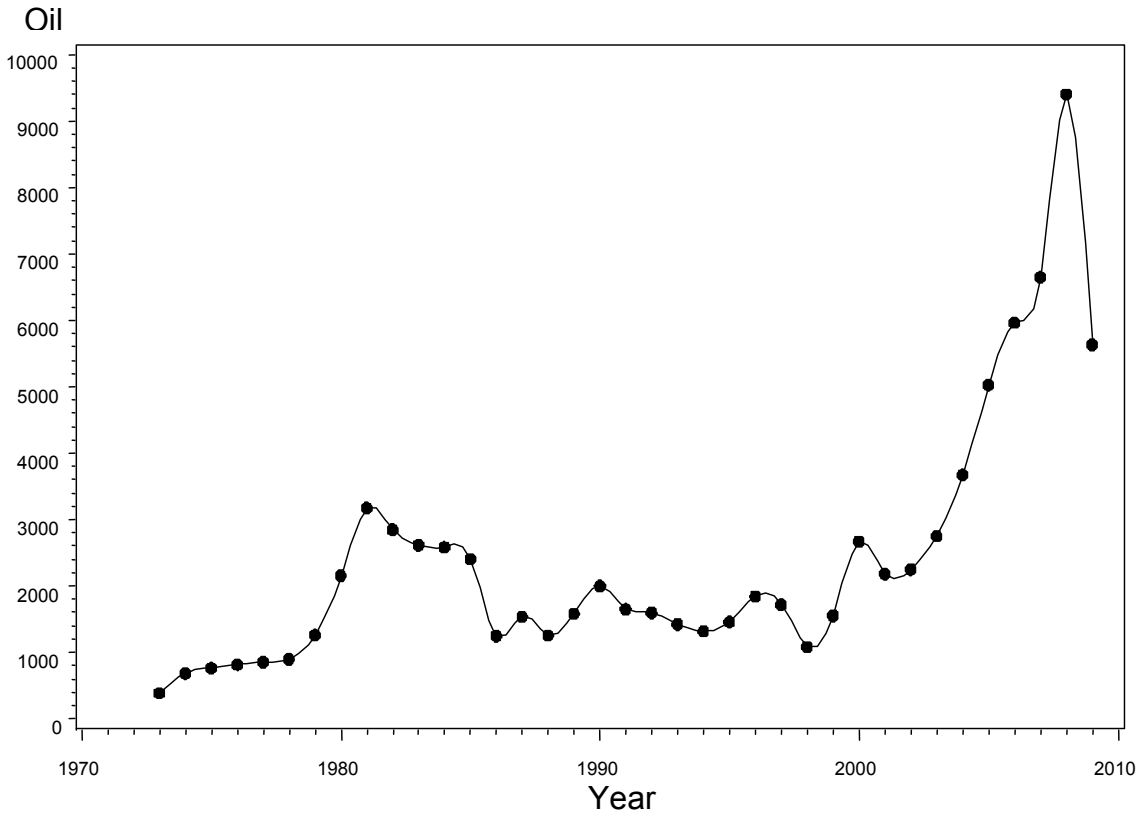
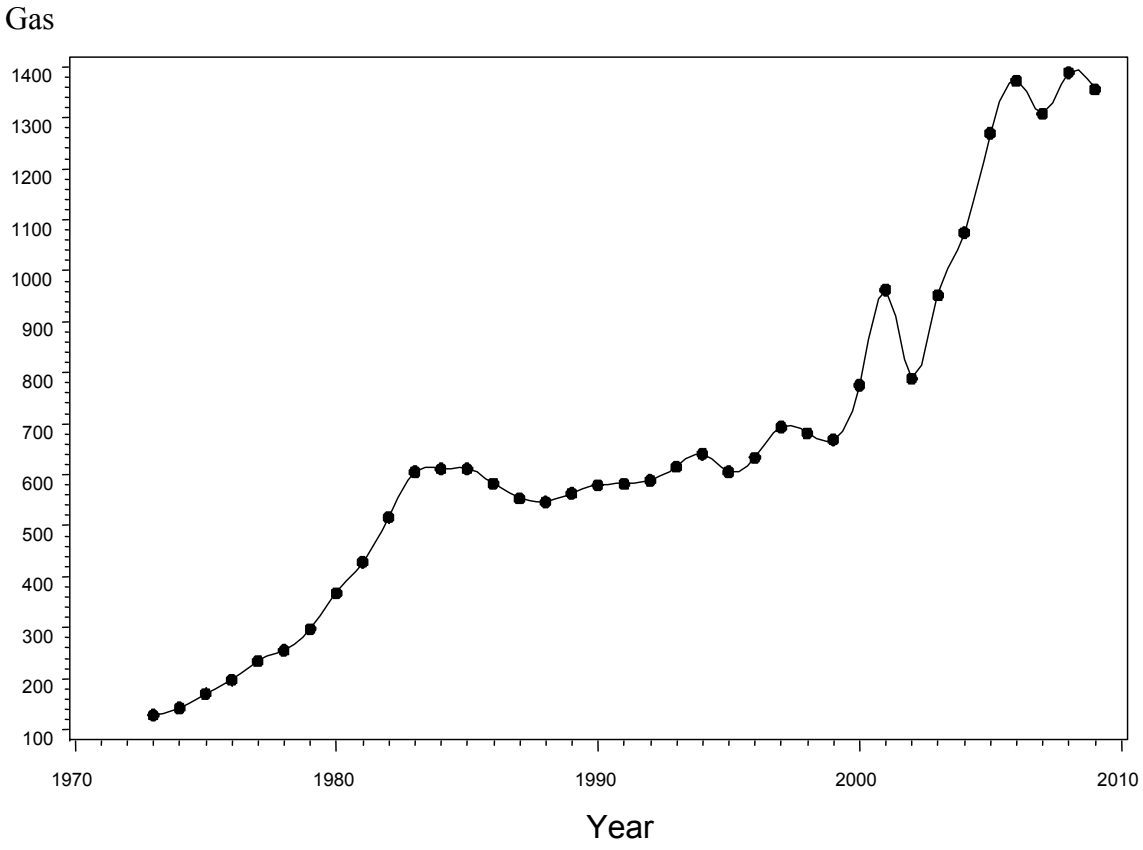


Figure 3: Plot of natural gas price over years.



ECONOMIC FACTORS IMPACTING THE HOUSTON APARTMENT MARKET

Michael E. Hanna, University of Houston-Clear Lake
Stephen C. Caples, McNeese State University
Charles A. Smith, University of Houston – Downtown

ABSTRACT

This paper explores the demand and pricing of apartments in Houston, Texas from 1980 to 2011. Both the price of apartments and the occupancy rate during this period are considered. The purpose of this paper is to determine how factors such as employment and population affect the demand for apartments. Some simple models are developed to forecast future occupancy rates and rental prices.

INTRODUCTION

The objective of this paper is to develop models to help forecast occupancy rates and rental rates for apartments in the Houston area. This would help investors understand the apartment market and make better decisions about new apartment construction. The price of rental units is decided by supply and demand for those units just as it is with other goods. The actual relationships among factors impacting supply and demand are the focus of this paper.

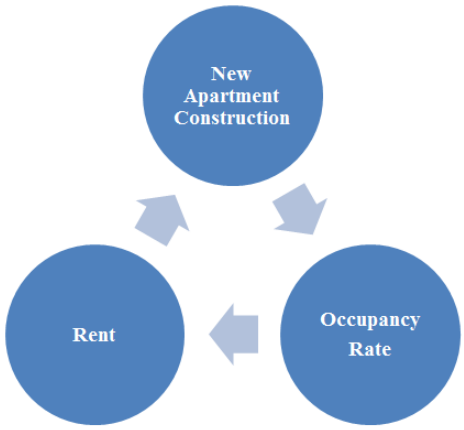
BACKGROUND AND LITERATURE REVIEW

Estimating future price for apartments has historically presented problems to practitioners in the field. Supply and demand are constantly at work in the market place. Figure 1 provides a simple model for the relationship between rental rates, new apartment construction, and the occupancy rate for apartments. It is circular because as units are filled and occupancy rate increases, there is an incentive to raise the rates because the demand is increasing relative to the supply of units. As the rents increases, developers have an incentive to build more units. As more units become available, the occupancy rates are again affected but in the opposite direction.

However, there are other factors external to the model shown in Figure 1 that must be considered. Two important ones are employment and population. If population is increasing faster than apartment availability is increasing, occupancy rate may continue to increase as demand outpaces the increase in supply. This may lead to another increase in rents, and the process continues. What might cause population to increase? One cause would be a growth in

jobs (employment). Even with these factors considered, this is a simple model. Other factors, such as changes in the supply due to older apartments being torn down, new single family rentals on the market, may impact this process. Further study in this area is ongoing.

Figure 1. The Apartment Rental Cycle



Understanding the demand for apartments and other commercial space has been the focus of many studies. Malizia (1991) recognized that long-term demand-side forecasting models needed to include economic development variables. Wheaton and Torto (1990) linked job growth to industrial supply and demand. There is a plethora of empirical data linking employment to various factors influencing demand for real estate, Numerous methods and models to forecast one aspect of real estate or another have appeared. In one of these, Valente, Wu, Gelfand and Sirmans' (2005) present a spatial model for predicting apartment rents.

Lentz and Tse (1999) present models to relate the performance and needs of the goods market to the demand for commercial real estate. They noted, "The commercial real estate market is frequently observed to be in an extended state of disequilibrium." Since there is a time lag between the beginning of the construction cycle and the time when the finished space is available for rent, it can be difficult to make an accurate estimate of future space demands. It is common to overbuild or fail to build enough space simply because the market changed at some point during the construction cycle. The decision to build new apartment space should be made after weighing expectations of future demand, apartment space under construction, absorption rates and the amount of vacant space already in the market. Lentz and Tse further observed," With future demand uncertain, the supply (quantity) of space and the realized demand for space may not match. If the supply is less than the realized demand, the space producer will be able to lease out all the new space. On the other hand, if the supply is greater than the realized demand, the excess supply will cost the space producer holding costs on the vacant units."

THE HOUSTON MARKET

Figure 2 provides information about the mean rental rates for apartments from 1980 to 2011. In the early 1980s, the rents were rising and reached a peak in 1982. They began falling and did not return to the 1982 high until 1990. During this same period, the occupancy rate (percentage of apartments that were rented) demonstrated a similar pattern as seen in Figure 3. With more vacant apartments, the apartment managers were forced to lower the rents and provide other incentives to attract new tenants and to keep the existing ones.

Figure 2. Mean Rental Rates for Apartments in the Houston PMSA from 1980 to 2011

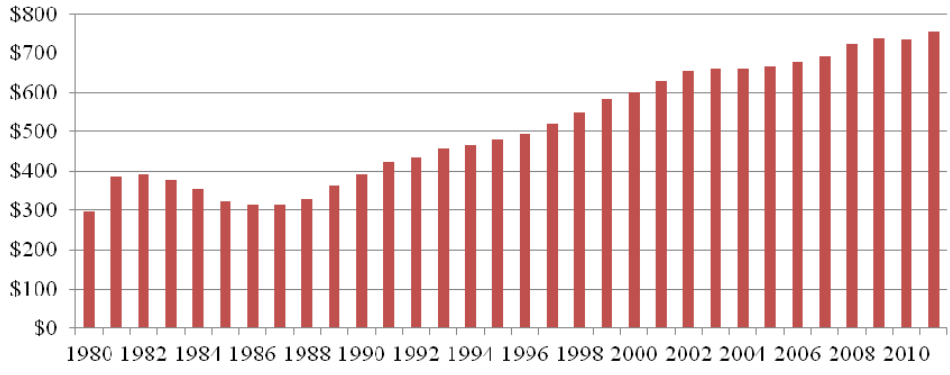
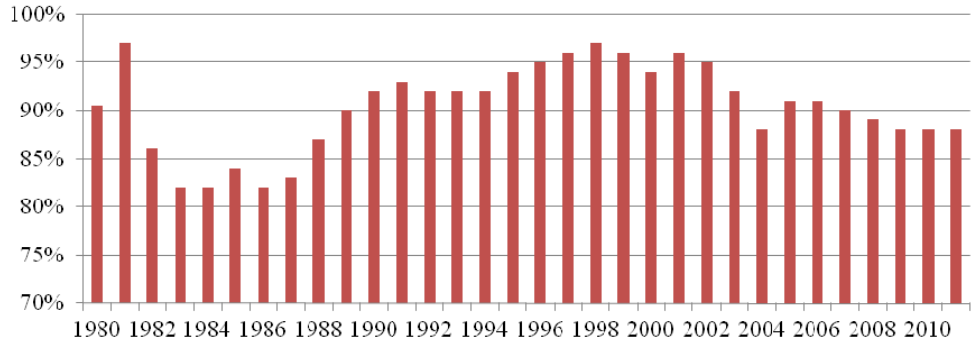
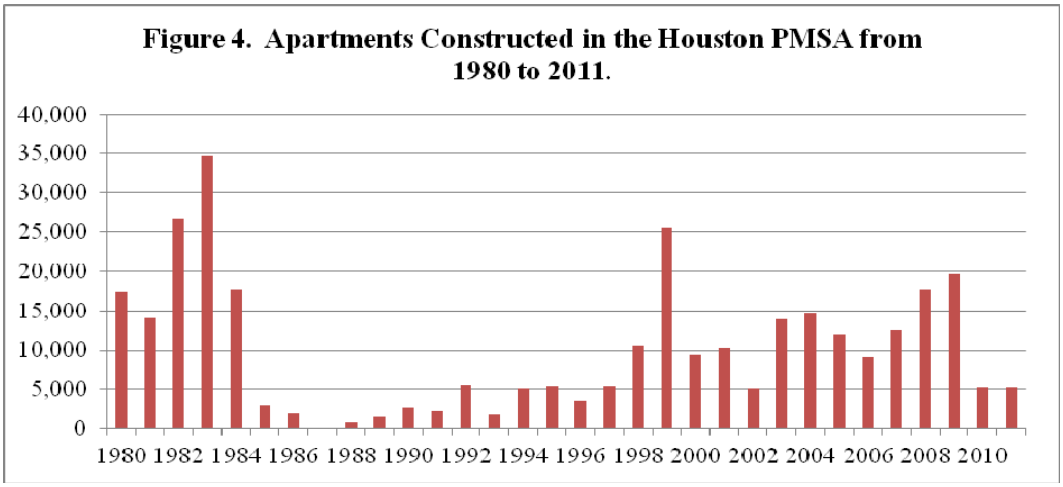


Figure 3. Occupancy Rate for Apartments in the Houston PMSA from 1980 to 2011.

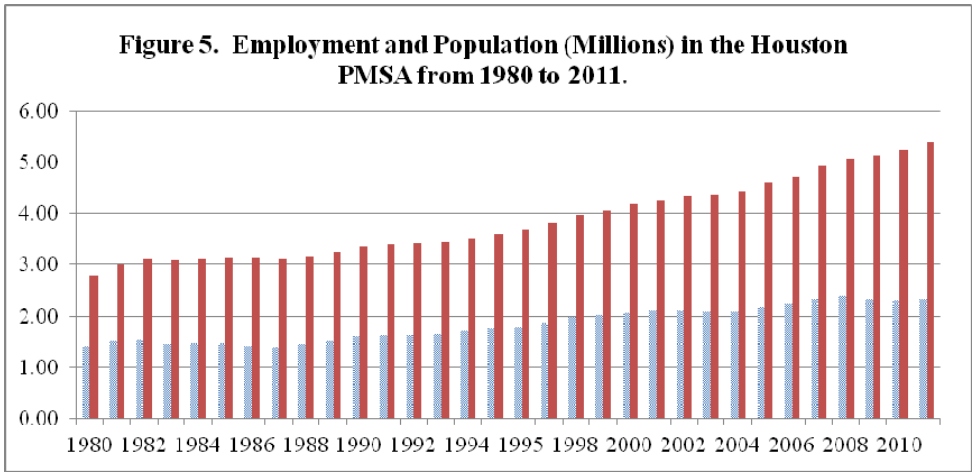


With lower rents and lower occupancy numbers in the mid-1980s, there was little need to construct new apartments. Figure 4 shows the construction of new apartments, and the early 1980s saw a large number of units coming onto the market. This was consistent with the increasing rents and occupancy rates at that time. What caused this? In the prior decade, the 1970's, Houston's economy was one of the fastest growing in the nation, mainly due to

Houston's role in the petroleum industry and the jobs created related to that industry. During this time, Houston construction was consistently among the highest in the nation.



On the demand side, it was the strong Houston economy that attracted more and more people to the Houston area looking for employment. Figure 5 shows how these two paralleled one another during the 1980s. The growth in population from the 1970s and early 1980s caused an increase in the demand for housing, which caused rents and occupancy of apartments to rise. Houston was very dependent on the petroleum industry in the 70's and 80's, and it obviously prospered when oil prices went up but it suffered as oil prices went down (e.g. mid-1980s), resulting in tremendous fluctuations. The Houston area is more diversified today, but the petroleum industry is still a large factor in its economic well-being.



It should be noted that, while increase in population may be due to employment opportunities, it also may be caused by other things. For example many retirees have chosen to move to Florida not because of employment, but because of the weather or lifestyle or other factors. However, in Houston, the growth in population was strongly tied to job opportunities as the correlation between jobs and population was about 98 percent. Thus, in this study, our focus is on employment rather than population.

As we have seen, the changes in employment caused changes in the demand and in the rental prices for apartments. However, alternative types of housing in a particular market can also affect the apartment rents. For example, if there is a glut of single family houses to rent and the rents are relatively low, then apartment rents will be curtailed to some degree. This is also true of duplexes and other types of available rental housing.

In addition to single-family rental properties, the price and availability of houses for sale also may impact the demand and the rents for apartments. Many people rent as an alternative to buying a home or they may rent while saving for a down payment or looking for a house to buy. Factors influencing home purchases may be local in nature, such as availability and pricing of alternative housing, but they may be national in nature, such as interest rates. As interest rates go down homes become more affordable and the impact on apartment rents is negative, since people are buying and don't need to rent. As interest rates go up, as they did in the 1970's, they make buying a house much more difficult. Higher interest rates result in higher monthly payments, which is often the determining factor in whether a person buys a new house or chooses to rent. Government regulations and incentives also affect home buying which in turn affects apartment rental rates. Building codes and zoning laws are a few examples that most people are familiar with (these tend to be local), but changes in tax rules have a large impact as well (these are mostly national in scope). In the 1970's the government gave a tax credit up to 5% in some cases to people who bought new homes. Recently tax credits of up to $8,000 were made available to some home buyers. This greatly reduced the price and impact of things like down payments on houses and stimulated the sale of houses. The new home purchases had a negative impact on the demand for rental units, and this lowered the rents that apartment owners could charge.

While many things impact the rents and occupancy rates for apartments, this study will be limited in scope. Housing rentals and purchase are not considered in this study, but they will be the focus of future research.

THE ANALYSIS

In trying to understand the Houston apartment market, we focus on two variables – monthly rent and occupancy rate (percentage of apartments that are occupied). The data used in our analysis included employment, population, and apartment units constructed. Appropriate

lags on some variables were tested, and the growth in the number of jobs and growth in the population were computed when appropriate.

A regression model was developed to predict the average monthly rent adjusted for inflation. The independent variables used in the stepwise regression model were occupancy rate, new apartments constructed, change in population and change in population lagged 1 year, and change in employment and change in employment lagged 1 year. The only two variables to enter the regression model were new apartments constructed and occupancy rate. Together they provided an r-squared value of 0.70. Certainly we expect the occupancy rate to be highly correlated with the rent. As an apartment approaches 100% occupancy, the rents would typically be raised, although this might take some time due to lease conditions.

The next model investigated the factors impacting the occupancy rate of the apartment. To see if it would be possible to predict changes in occupancy rates, a stepwise regression model was used. The variables included in this model were change in employment, change in population, and apartments constructed. The same three variables were lagged one year and also included in the model. Upon running the model, the employment change was the first variable to enter the model, and this was followed by the employment change lagged one year. The r-squared for the resulting regression equation was 0.36. This increase in occupancy rate that occurs when employment increases is certainly to be expected. People move in to the area to take a new job, and an apartment is one housing alternative that is popular.

The final model that we developed was intended to see how developers might use this data to make decisions about apartment construction. The dependent variable in the model was number of units constructed. These new units do not appear overnight as the construction process is often quite long. Therefore, the variables used to predict this are all lagged one year. These are lagged variables for employment, population, adjusted rent, and occupancy rate. The significant variables found using the stepwise regression were the adjusted rent lagged one year and the adjusted population lagged one year. The resulting r-squared value was 0.66.

SUMMARY AND CONCLUSIONS

The data from the Houston PMSA for the years 1980 to 2011 were used to develop models to investigate the apartment market. The best regression model to predict the average rent included the new apartments constructed and the occupancy rate, and it had an r^2 of 0.70. A high occupancy rate would drive up the price. The fact that new apartments were constructed would indicate that there was a high demand. While other factors impacted the new apartment construction, this variable captures the information provided by those other factors.

The best regression model for predicting the occupancy rate included the change in employment and the change in employment lagged one year. This is consistent with what was expected. However, the model yielded an r^2 of 0.36, and this was somewhat lower than was expected.

The final regression model was used to predicted new apartment construction. The two significant variables in this model were the adjusted rent lagged one year, and the population change lagged one year. Higher rents would make building new apartments more profitable, and more people moving into the area would cause an increase in the demand for apartments. The model yielded an r^2 of 0.66.

The models presented here provide some information that may help us understand the Houston apartment market. However, there is much more to be discerned. Future studies in this area will incorporate additional variables, including alternative housing, in an attempt to develop even better models and to better understand this market.

REFERENCES

Benjamin, John D., Glenn W. Boyle, and C. F. Sirmans (1990). Retail Leasing: The Determinants of Shopping Center Rents. *AREUEA Journal* 18(3), 302-312.

Hanna, Michael E., Stephen C. Caples, Charles A. Smith, and Charles P. Rollins (2007). "Factors Impacting Price for Retail Space in Houston". *Economics and Economic Education Research Journal 8*(2), 93-103.

Lentz, George H. and K.S. Maurice Tse (1999). Supply Adjustments to Demand Shocks in the Commercial Real Estate Market. *Real Estate Economics* 27, 231-262.

Malizia, Emil E. (1991). Forecasting Demand for Commercial Real Estate Based on the Economic Fundamentals of U. S. Metro Markets. *The Journal of Real Estate Research* 6(3), 251-265.

Valente, James, ShanShan Wu, Alan Gelfand, and C.F. Sirmans (2005). Apartment Rent Prediction Using Spatial Modeling. *The Journal of Real Estate Research* 27(1), 105-136.

Wheaton, William C. and Raymond G. Torto (1990). An Investment Model of the Demand and Supply For Industrial Real Estate. *AREUEA Journal* 18(4), 530-547.

UNSTABLE RELATIONSHIP BETWEEN THE FED'S MONETARY POLICY ACTIONS AND THE U.S. STOCK MARKET

Anthony Yanxiang Gu, State University of New York at Geneseo
Xiaohui Gao, Shanghai University of Finance and Economics

ABSTRACT

The stock market and target federal funds rate is positively related over the long run although the relationship is not stable with unstable time lags. This is consistent with the fact that the Fed usually raises the target interest rate when economic growth is strong and the stock market is in bullish trend, and cuts the interest rate when growth is declining and stock market is in bearish trend. The relationship is apparently negative between the two variables when the stock market was set lagging the interest rate by three years, which may imply an obvious time lag between the two variables. Granger causality tests show that the Fed raises interest rate after the stock market has been rising and cuts interest rate after the stock market has been declining. The relationship between monetary aggregate M1 and stock market is negative, and that between M2 and stock market is positive.

Keywords: monetary policy action, stock market trend, time lags
JEL classification: E44 E52, G1

INTRODUCTION

Economists have been trying to find and explain stock market's respond to the Fed's monetary policy actions. Kuttner (2001) and Bernanke and Kuttner (2005) find that investors react to surprise policy actions. Brown and Cliff (2005), Backer and Wurgler (2006, 2007), Kumar and Lee (2006) and Kurov (2010) show that investors' sentiment play a major role in stock market's reaction to monetary policy actions. Chen and Gu (2011) find that stock market declines immediately after regular Fed announcement on most of the announcement days when change in the target federal funds rate exactly meets the consensus estimate. These researches focus on short-run response of stock market to Fed action. Chen (2007), Basistha and Kurov (2008) and Kurov (2010) find that monetary policy actions have stronger impact on stock market in bearish market trends and in recessions.

However, the relationship between the Federal Reserve's monetary policy actions and stock market is not stable. The goals of the Federal Reserve System include economic growth, low unemployment rate, and price stability. The Fed's policy actions for these goals, such as

changing money supply and target interest rate of the federal funds, have shown apparent impacts on growth of the economy, inflation rate and unemployment rate. These impacts indirectly affect companies' profit and thus their stock prices. However, the effect of Fed action on these goal variables has time lags, e.g., Mishikin (2010) reports that Fed actions affect the economy with time lags of a year or longer, which makes measuring the effect even more difficult. The stock market usually leads economic cycle by several months. Although condition of the stock market is not explicitly a concern of the Fed, as Alan Greenspan said at a Congress hearing in 1999 that the stock market is not a consideration for the Fed's monetary policy even after he had said "over exuberance" about the stock market bobble in 1997, the Fed with Bernanke being the chairman may actually pay attention to stock market condition as increasingly more economists, including these at the Fed, have recognized the wealth effect of the stock market on consumer behavior and on the economy.

During the great bullish market period in the 1990s, The Fed raised the federal funds target rate from 1994 to June 1995, then reduced it moderately and kept the rate at around 5.25 percent until early 2000, even after the stock market crashed and started the great bearish period in March 2000, the Fed raised the rate to 6.5 percent and kept it through the rest of the year. Nine months later, the stock market had declined significantly, then the Fed started cutting the federal funds target rate in January 2001.

In this study we focus on exploring the long-run relationship between the federal funds target rate and the stock market from 1994 to March 2011, we also examine the relationship between monetary aggregates M1 and M2 and the stock market in the period although former Board of Governors chairman Alan Greenspan testified in congress that the Fed would no longer use any monetary aggregates as a guide for conducting monetary policy in January 1994. The purpose of this study is to find how close and with what time lags the Fed's policy actions and the stock market are related or the impact of the policy action on the stock market. The results of this study should provide investors with useful information for developing investment strategies.

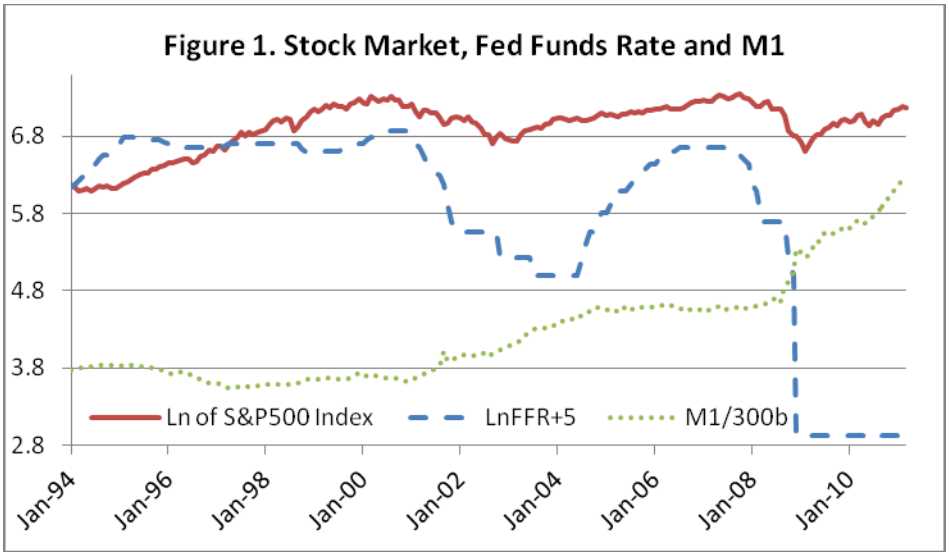
DATA

Data of the federal funds target rate and M1 and M2 is from the Feral Reserve Bank of New York. Data of the S&P 500 monthly Index is from Finance.Yahoo.com. Most of the times the Fed changes its federal funds rate target around the middle of the month, and occasionally change the rate twice in a month, so we calculate the time weighted average of the target rates and use it for the month.

THE ANALYSES

Generally over economic cycles, the stock market is growing when the economy is growing strongly, and the Fed raises the federal funds target rate; the stock market is declining

when the economy is in a recession, and the Fed cuts the interest rate. Obviously there is some time lag among the three variables although length of the time lag is not stable and hardly predictable. Thus, we may expect that the federal funds target rate is positively related to the S&P 500 Index. We display the three time series in Figure 1. In order to make the comparison more apparent, we adjusted the scale of the federal funds rate, the S&P 500 Index and M1, i.e., we took the natural logarithm of the index and of the federal funds rate, then add 5 to the logarithm of the federal funds rate, and adjust the scale of M1 by dividing the quantity with 300 billion. As depicted in Figure 1, one can see the relationship between the S&P 500 Index and the federal funds target rate is not stable and changes direction over time. This may indicate that the impact of the Fed's actions of federal funds on the economy hence the stock market is not simultaneous and the relationship is far from being perfect. Movements of the federal funds target rate and the S&P 500 Index are roughly in the same direction most of the time with obvious time lag. Movements of M1 and the stock index are not in the same direction most of the time although both time series are increasing in the long-run. Obviously the M1 aggregate increased much more than the stock index for the sample time period, and growth rate of M1 fluctuates significantly. Specifically, M1 declined and was relatively flat while the stock market was rising significantly from 1994 to March 2000, M1 increased but the stock index declined from April 2000 to September 2002, and M1 increased sharply but the stock market dropped sharply from 2008 to early 2009. M2 is basically mono-increasing from $3480 billion to $9112 billion over the sample time period, it can be expected to have a positive relationship with the stock market over the long-run, so we exclude it in the figure.



We conduct regression tests to reveal the relationship between the stock market and the Federal Reserve's policy action variables, namely, the target federal funds rate, M1 and M2.

$$S = \alpha + \beta_1 FR ++ \beta_2 M_1 + \beta_3 M_2 + \varepsilon \quad (1)$$

where,
S = natural logarithm of the S&P 500 monthly Index
FR = the federal funds rate
M_1 and M_2 are the monetary aggregates, and
ε is the error term

The regression results are reported in Table 1. As shown in the table, the S&P 500 Index is significantly positively related with the federal funds rate, this relationship is obvious in Figure 1 for most of the time in the sample period, particularly from 1999 to 2008. The relationship between movements of the stock market and M1 is significantly negative, which can also be observed in Figure 1 for most of the time. The relationship between the index and M2 is significantly positive as M2 is almost mono-increasing and the stock market was moving upward much longer than moving downward in the period.

Table 1: REGRESSION OF THE S&P 500 INDEX ON FEDERAL FUNDS RATE AND MONETARY AGGREGATES

Time	Intercept	FR	M1	M2	adj. R^2
t_0	1043.658	70.274	-1.716	0.343	0.693
	(7.281)***	(8.687)***	(-11.159)***	(19.358)***	
t_{36}	31.698	-3.909	-	-	0.154
	(6.513)***	(-5.659)***	-	-	

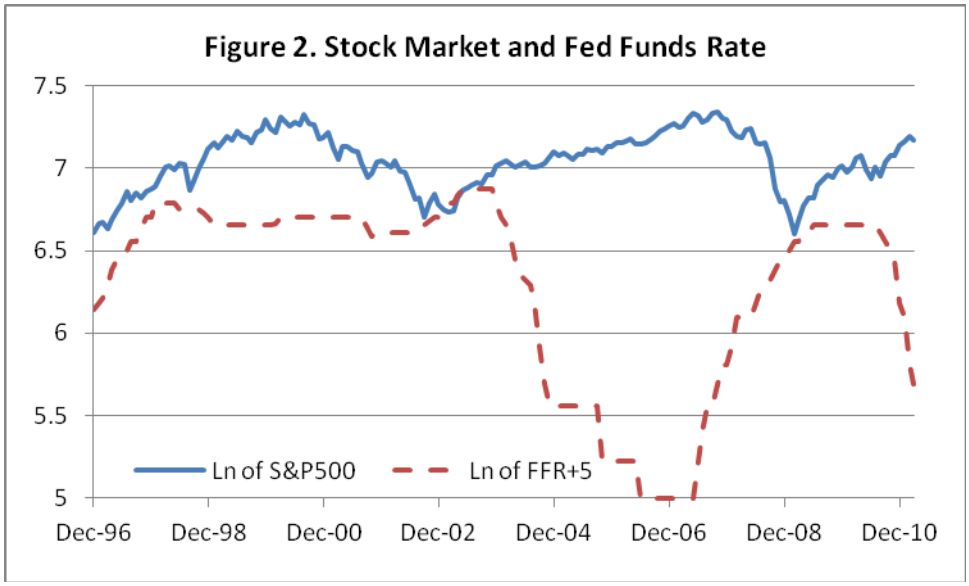
In order to find at what time lag the relationship between the stock market and the federal funds target rate is the most significant, we have conducted a series of regression tests,

$$S = \alpha + \beta FR_{t-i} + \varepsilon \quad (2)$$

where,
S = natural logarithm of the S&P 500 monthly Index
FR_{t-i} = the federal funds rate, i = 0…40
ε is the error term

The regression with t=0 results in a coefficient of -0.018, and a t-value of 1.033. The relationship is almost zero and insignificant, this result is because the relationship is not stable, there is uncertain time lag with the impact of Fed action and the economy, and regression analysis provides a coefficient that is the mean of positive and negative relations. The regression

results show the highest R-squire, adjusted R-squire, and absolute t-value with i = 36. As shown in Table 1, the coefficient is significantly negative, which indicates that the federal funds target rate is the most significantly negatively related to the 36-month-later S&P 500 Index. This pattern is apparent from 1998 as shown in Figure 2 where one can see that the federal funds target rate was declining from July 1995 to November 1998 and then increasing until December 2000, lagging roughly 36 months, the S&P 500 Index was increasing and then declining from late 1997 to late 2001. The negative relationship is even more apparent later. The federal funds target rate was declining from January 2001 and then rising from July 2004, and then declining from September 2007, three years later, the S&P 500 Index was rising and declining from October 2002 to early March 2009, and then rising from March 2009 to April 2011. This pattern implies that the stock market may enter into a bearish trend two to three years after the Fed started raising the federal funds rate in a cycle, and the stock market may enter into a bullish trend two years after the Fed started cutting the interest rate. However, there is no doubt that the time lag between the movements of the two variables is not stable.



The relationship between federal funds target rate and stock market as shown in the figures and as revealed in the regression analyses does not explain whether changes in the federal funds target rate cause stock market movement, or stock market trend leads Fed actions. To further explore the relationship we conduct Granger causality tests. Granger (1969) proposed a test to determine whether or not a series x_t "causes" changes in the series y_t. A critical implication of Granger causality tests is that they do not prove causality in the general sense; rather they illustrate Granger-causality. That is the use of Granger causality tests reveal whether or not current and or lagged values in the series x_t improve our ability to forecast changes in y_t.

The standard bi-variate Granger causality test is based on OLS regressions of the following two equations:

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \ldots + \alpha_p y_{t-p} + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \ldots + \beta_p x_{t-p} + u_t \quad (3)$$

$$x_t = \alpha_0 + \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \ldots + \alpha_p x_{t-p} + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \ldots + \beta_p y_{t-p} + u_t \quad (4)$$

The test for whether or not *x* Granger causes *y* is based on the results of an F-test on the joint hypothesis:

$$H_0 \qquad \beta_1 = \beta_2 = \cdots = \beta_p = 0.$$

If the null hypothesis is rejected, we conclude that x Granger-causes y. The same test is also applied to test whether y Granger-causes x. From the above regressions there are four potential outcomes, i.e. x Granger-causes y, y Granger-causes x, or causality runs in both directions, and finally neither y Granger-causes x or vice versa. Perhaps the most useful results would be where causality runs in only one direction, which would imply for example that by knowing past values of x, the forecasts of y are improved.

There are two potential issues to be considered when using Granger-causality tests. First, the data for both series must be transformed to eliminate autocorrelation, for example, first differencing of each series. Second, the choice of lag length may play an important role. Generally, longer lag lengths are preferable to allow for the effect of all relevant past information on x (y) to have an effect on y (x). In our specific case the use of p = 10 includes the information from two full business weeks.

Results of the Granger-causality tests are reported in Tables 2, in order to save space we report results for 12-month lag only. The tests reveal that the S&P 500 Index Granger causes the target federal funds rate while the target federal funds rate does not Granger cause the S&P 500 Index. This indicates that the Fed changes its target federal funds rate following the stock market. As discussed above, the Fed raises the interest rate when growth of the U.S. economy is strong and the stock market is already in an upward trend, and cut the interest rate when growth of the economy is weak or the economy declines and the stock market is already in a downward trend.

Table 2: RESULTS OF PAIRWISE GRANGER CAUSALITY TEST

Hypothesis	Lag	F-Statistic	Prob.	Result
FR does not Granger Cause S	12	0.97537	0.4743	Not reject
S does not Granger Cause FR	12	2.82102	0.0015	Reject

CONCLUSION

In this study we examine the relationship between the Federal Reserve's monetary policy actions and the U.S. stock market. We reveal that the federal funds target rate and the S&P 500 Index is positively related over the long run, but the relationship is not stable and with unstable time lags. This corresponds to the fact that the Fed raises the target interest rate when the U.S. economic growth is strong while the stock market is in a bullish trend and cuts the interest rate when the U.S. economic growth is declining while the stock market is in a bearish trend. There is an apparent negative relationship between the two variables when the stock market is set lagging the target interest rate by three years, this may imply an obvious time lag between the two variables. The implication for investors is that the stock market may enter into a bearish trend two to three years after the Fed started raising the federal funds rate in a cycle, and the stock market may enter into a bullish trend two years after the Fed started cutting the interest rate. The Granger causality tests reveal that the Fed's actions on the federal funds rate follow the stock market. The relationship between monetary aggregate M1 and the stock market is negative over the long run, and that between M2 and the stock market is positive because M2 is almost mono-increasing in the sample period. Results of this study imply that investors may consider Fed's actions on the federal funds target rate when developing their investment strategies, i.e., prepare to see a bearish market when the Fed has been raising interest rates for two years and prepare to ride on a bullish market when the Fed has been cutting interest rate for two years.

REFERENCES

Aziza, F.O.F. (2011). The effects of monetary policy on stock market performance: A cross-country analysis, *SSRN Working Paper Series,* posted on January 23, 2011, from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1743834

Baker, M. & J. Wurgler (2006). Investor sentiment and the cross-section of stock returns. *Journal of Finance,* 61(4), 1645-1680.

Baker, M. & J. Wurgler (2007). Investor sentiment in the stock market. *Journal of Economic Perspectives,* 21(2), 129-151.

Basistha, A. & A. Kurov (2008). Macroeconomic cycles and the stock market's reaction to monetary policy. *Journal of Banking and Finance,* 32(12), 2606-2616.

Bernanke, B.S. & K.N. Kuttner (2005). What explains the stock market's reaction to federal reserve policy? *Journal of Finance,* 60(3), 1221-1257.

Brown, G.W. & M.T. Cliff (2005). Investor sentiment and asset valuation. *Journal of Business,* 78(2), 405-440.

Chen, S.S. (2007). Does monetary policy have asymmetric effects on stock returns? *Journal of Money Credit and Banking,* 39(2-3), 667-688.

Chen, S. & A.Y. Gu (March 2011). Expectation fulfillment effect on the Fed's announcement day: Evidence from the U.S. stock market. Presented at the *Midwest Economics Association Annual Meeting,* Saint Louis, Missouri, United States.

Granger, C.W.J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica,* 37(3), 424-438.

Kumar, A. & C.M.C. Lee (2006). Retail investor sentiment and return co-movements. *Journal of Finance,* 61(5), 2451-2486.

Kurov, A. (2010). Investor sentiment and the stock market's reaction to monetary policy. *Journal of Banking and Finance,* 34(1), 139-149.

Kuttner, K.N. (2001). Monetary policy surprises and interest rates: Evidence from the Fed funds futures market. *Journal of Monetary Economics,* 47(3), 523-544.