Volume 15, 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., PO Box 1032, Weaverville, NC 28787, USA. Those interested in communicating with the *Journal,* should contact the Executive Director of the Allied Academies at info@.alliedacademies.org.

EDITORIAL REVIEW BOARD

EDITORIAL REVIEW BOARD

TABLE OF CONTENTS

LETTER FROM THE EDITOR

The *Journal of Economics and Economic Education Research* 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 *Journal* is the official publication of the Academy of Economics and Economic Education, which 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. The first category of manuscripts we desire is theoretical and empirical research which can advance the discipline of economics. The second category is research which can advance the effectiveness of economic education.

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 through the Allied Academies webpage at www.alliedacademies.org.

Grady Perdue
University of Houston-Clear Lake

Martin Milkman
Murray State University

John Marcis
Coastal Carolina University

STUDENT RESPONSE TO ANNUITY FORMULA DERIVATION

C. Alan Blaylock, Henderson State University

ABSTRACT

This paper presents an intuitive approach to deriving annuity formulas for classroom use and attempts to assess student response to such presentation in a principles of finance course by observing student behavior and surveying student sentiment. After presenting three derivations of an annuity formula, a quiz was administered to determine the correct use of a particular derivation of the student's choosing. The derivation used and its correct use by students is reported. In addition, a survey is conducted to assess student sentiment. Implications for instructors are also discussed.

INTRODUCTION AND LITERATURE REVIEW

This paper attempts to assess student reaction to teaching the derivation of annuity formulas in a principles of finance course. The research methods include observing student behavior and surveying student sentiment. The results presented in this paper allow the instructor some insight into student perceptions of how annuity formulas should be taught.

Many students find the use of mathematics in any discipline somewhat difficult. This is indeed true for finance courses as well, especially for the finance principles course that is usually required of all business majors. Some instructors allow the use of formula sheets during testing. For those students that are not given such luxury, memorizing and comprehending mathematical formulas can be a daunting task. Many students distinguish between material that is more "concept" and that which requires the use of formulas. They fail to realize that the two are inseparable; formulas are simply concepts written in mathematical terms. If students were allowed to see the reasoning behind the formulas, they may better be able to understand and use those formulas.

Time Value of Money (TVM), the valuing of cash flows through time, is a major topic in the teaching of finance. Of the various formulas available to adjust dollar values, those involving the present and future value of annuities may be regarded as perhaps the most difficult to comprehend in the principles course. This may be due to the fact that they are lengthy and their logical derivation is not so easily grasped in the way they are usually presented. Most corporate finance texts write the Present Value of an Annuity (PVA) formula using a constant rate for each payment period in the form

$$\text{PVA} = \text{PMT}\left[\frac{1 - \frac{1}{(1+\text{r})^{\text{n}}}}{\text{r}}\right] \quad (1)$$

where PVA is the value of the annuity one period before the first payment, PMT is the annuity payment at the end of each period, r is the rate of change in dollar values per period, and n is the number of periods. Some other finance texts use other forms of the formula, some of which will be developed shortly. A common form of the Future Value of an Annuity formula is

$$\text{FVA} = \text{PMT}\left[\frac{(1+\text{r})^{\text{n}} - 1}{\text{r}}\right] \quad (2)$$

To a student with a limited math background, these formulas can look overwhelming in their concept and size. In an age without computers Sherritt (1944) suggests limiting the number of formulas, restricting the number of symbols across all formulas, and not requiring that formulas be developed. To further assist student learning he graphically connects each formula to a time line diagram.

In order to avoid referring to a geometric progression to explain annuity formulas Watson (1936) instead points out that the FVA formula is the compound interest of $1 divided by the interest rate per period and that the PVA formula is the compound discount on $1 divided by the interest rate per period.

In order to help students match a set of cash flows with an appropriate valuation formula Skinner (1994) begins with a geometric progression and derives a general "source" formula that can, by imposing different assumptions related to a set of cash flows, be used to calculate the present value of

an ordinary annuity
a perpetuity
a growing perpetuity
a growing annuity
a growing annuity with constant rates of growth.

Using this formula with varying sets of assumptions, "establishes the critical link between the structure of the cash flow to be valued and the appropriate model to be used" (Skinner, 1994, p. 87).

Most introductory finance texts simply omit any explanation of the annuity formulas. A few, however, do present to varying degrees the concept that an annuity is a piece of a perpetuity, and then use that concept to derive the annuity formulas. Notable texts that offer such presentations either in the main text, footnotes, or as web extensions or chapters include Brealey, Myers, and Allen (2014), Brigham and Ehrhardt (2014), Ross, Westerfield, and Jaffe (2013), and Welch (2009). Although seemingly overwhelming to a student in introductory finance the formulas actually represent this quite simple concept of valuing an annuity as a part of a perpetuity.

The connection an annuity has with a perpetuity is not easily seen in the often used forms of annuity formulas. However, assuming that students easily accept the perpetuity formula they may be better prepared to accept annuity formulas if they see that such formulas can derived from the simple perpetuity formula. Newcombe et al. (2009) cite several studies in which elementary school children were better able to determine solutions to fraction problems when they were able to connect such problems to "related informal experiences." Newcombe et al. also cite research that shows that confidence in one's math ability increases performance in science, technology, engineering, and mathematics (STEM) learning activities. Beginning with the simple perpetuity formula may provide students with familiar material that can be used to build confidence to use the more difficult annuity formulas.

The major question this paper asks is, "How do students respond to a presentation of the derivation of an annuity formula from the perpetuity formula?" This paper uses the method of observing student behavior and surveying student sentiment. First, students are presented with the derivation of the annuity formulas that is similar to Brealey, Myers, and Allen (2014), Brigham and Ehrhardt (2014), Ross, Westerfield, and Jaffe (2013), and Welch (2009). Three common forms of each of the formulas are given, with each form representing regressive "closeness" to the simple concept from which it was derived. As will be shown shortly, the first form is written in a way that better reflects the concept of its derivation (the most intuitive form). The subsequent two forms, each a derivation of the previous form, progressively distance themselves from the initial simple derivation (less intuitive forms). After presenting the three forms, students are given a TVM quiz which requires them to write the formulas they used.

However, although some students may use one form, they still may have benefited from being exposed to other forms. This cannot be detected by observation. Therefore, students are asked if they did receive some benefit from being taught how the annuity formulas were derived. Both negative and positive reactions are anticipated. Although some students may benefit, others may deem such derivations too complex, deep, and confusing. By revealing the benefits and costs involved through student usage and perceptions, the results presented in this paper allow the TVM instructor some insight into how the annuity formulas should be taught. Such analysis of students' response to formula derivation is lacking in the literature.

DERIVING THE ANNUITY FORMULAS

Before the results are reported, the derivation of the PVA formula into three forms is given as a sample of what students were presented. One of the simplest formulas in finance is that of the Present Value of a Perpetuity (PVP). The PVP formula is

$$\mathrm{PVP} = \frac{\mathrm{PMT}}{\mathrm{r}} \tag{3}$$

where PMT is the perpetual periodic cash flow, and r is the discount percentage rate. The presentation presumes students understand this simple formula.

The simple concept in deriving the present value of an annuity results from the fact that the annuity is simply a part of a perpetuity; the only difference between a perpetuity and an annuity is that a perpetuity has an infinite stream of payments whereas an annuity has a finite stream of payments. The object of deriving the annuity formula from the perpetuity formula rests in

subtracting the part of the perpetuity that is not needed in order to leave the part of the perpetuity that is needed, the annuity. This is illustrated in Figure 1. Perpetuity B is subtracted from Perpetuity A to result in Annuity Z. Since the part of Perpetuity A that is not needed extends into infinity the subtracted part is, itself, is an annuity (Perpetuity B). The remaining payments of Perpetuity A (that no longer extend into infinity) would form the desired annuity, Annuity Z. Of course, in order to subtract the value of the future perpetuity from the value of the present perpetuity, the two values must be at the same time period. The two values must be adjusted to the desired time period, and the resulting difference would be the value of the annuity at that time period.

This process is illustrated in more detail in Figure 2 by deriving the present value of a four-period annuity (value at t = 0; n = 4). The value of a perpetuity at period 0 is calculated. The value of the fifth and remaining payments needs to be subtracted in order to retain and reveal the value of the four-period annuity. This is done by first calculating the value at period 4 (t + n) of those unwanted payments. Since the unwanted payments also form a perpetuity, this results in the value of that perpetuity at period 4. Since the value of the annuity is to be calculated at period 0, the value of the unwanted payments, PVP_4, is adjusted to period 0 using the simple present value formula (which would have been previously presented in class). The value at period 0 of the desired annuity can then be found by subtracting PVP_4 from PVP_0. Thus the Present Value of an Annuity (PVA) can be calculated as

$$\mathrm{PVA}_0 = \frac{\mathrm{PMT}}{\mathrm{r}} - \frac{\mathrm{PMT}}{\mathrm{r}}\left[\frac{1}{(1+\mathrm{r})^4}\right] \tag{4}$$

Or more generally

$$\mathrm{PVA}_\mathrm{t} = \frac{\mathrm{PMT}}{\mathrm{r}} - \frac{\mathrm{PMT}}{\mathrm{r}}\left[\frac{1}{(1+\mathrm{r})^\mathrm{n}}\right] \tag{5a}$$

where n is the number of payment periods proceeding (to the future of) time period t. This is the first form of the PVP formula and will be designated as Form A. The concept of how the

Figure 1
Deriving the Present Value of an Annuity

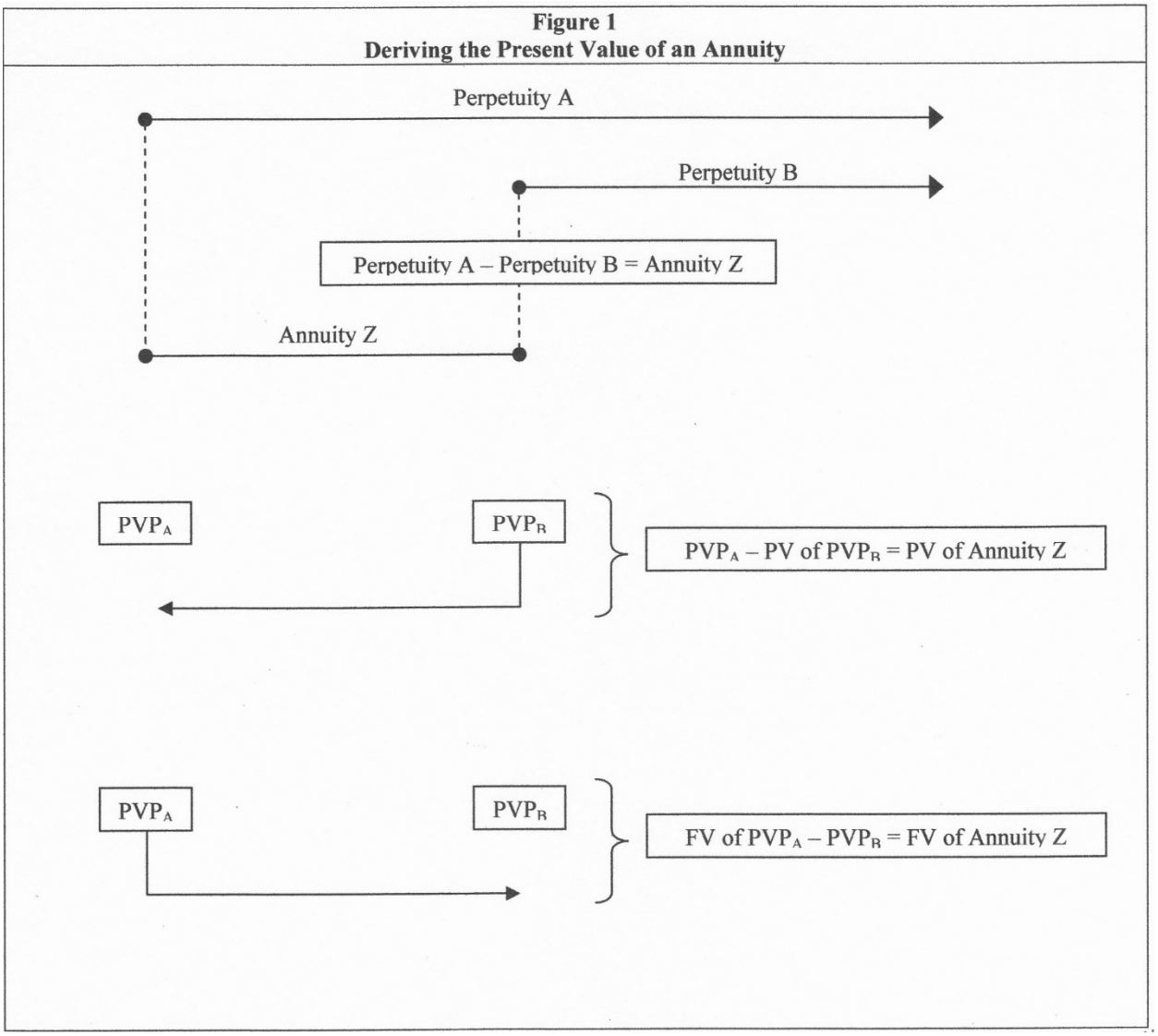
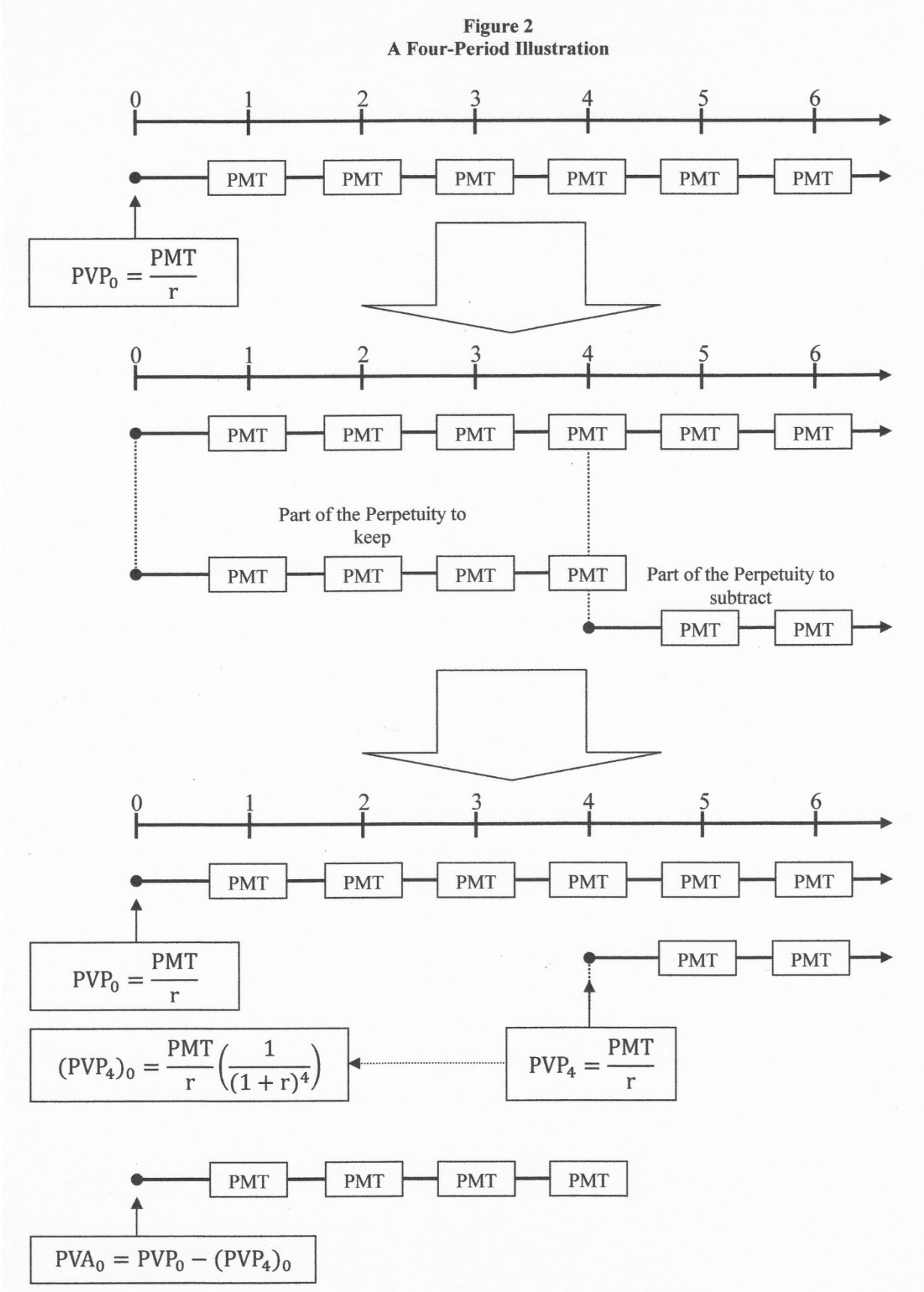


Figure 2
A Four-Period Illustration



annuity formula is derived is best seen in this form. Rearranging Form A results in forms B and C, respectively,

$$\text{FORM B: } PVA_t = \frac{PMT}{r}\left[1 - \left(\frac{1}{(1+r)^n}\right)\right] \tag{5b}$$

$$\text{FORM C: } PVA_t = PMT\left[\frac{1 - \frac{1}{(1+r)^n}}{r}\right] \tag{5c}$$

Notice that Equation 5c (Form C) is the same as Equation 1 presented earlier. A similar presentation is given to derive three forms of the Future Value of an Annuity (FVA) which are not presented here.

METHODOLOGY

Students were first presented the derivations of the future and present value annuity formulas. Students were not required to derive the annuity formulas at any time during the semester. Within a week of when the annuity formula derivations were taught, students were given a pop quiz to test their ability to calculate: (1) present value of an annuity (PVA), (2) future value of an annuity (FVA), (3) the payment of an annuity given its present value (PMT), (4) present value of a perpetuity (PVP). Having the students calculate the payment for an annuity is included to assess student reaction of having to manipulate any given form to solve for a missing variable. Although students could use a financial calculator, they were required to write the formula used in solving the problems.

Hypothesis tests are used to determine if any significant difference exists between the following:

1. The proportion of students who recalled a given formula and those who did not.
2. Of those who could recall a given formula, the proportion of students who recalled a particular form of that formula and those who recalled another form.
3. Of those who could recall a given formula, the proportion of students who calculated the correct answer and those who could not.
4. Of those who could recall a given form of a formula, the proportion of students who calculated the correct answer and recalled a particular form of a formula and those who calculated the correct answer and recalled another form.

Although students may use one of the three forms, they still may have benefited by being exposed to the two other forms. For instance, a student may use form C, but the student may be better able to use that form in calculating the correct answer because they were exposed to forms A and B and to their derivations. Therefore, a survey of student sentiment was conducted. Within the last two weeks of the semester, students were asked to (1) write the PVA and FVA formulas

in the form of their choosing together with the PVP formula, and (2) state in writing if exposure to the derivation of the annuity formulas benefited them in their classroom performance. Chi-square and related hypothesis tests are used to determine significance in regard to recalling a particular formula, a particular form of a formula, and if the student responded with a positive, negative, or neutral sentiment to the annuity formula derivations.

RESULTS

Table 1 presents the results of the pop quiz. A total of 86 students took the quiz. Of those students, a total of 77 (90%) were able to correctly write the PVP formula, more than the number who could correctly write the PVA or FVA formulas. This may be a somewhat obvious indication that the PVP formula is much easier to comprehend than the PVA or FVA formulas. Part B of Table 1 shows that of those who remembered the formulas, 90% actually correctly calculated the answer to the FVA problem compared to only 85% for the PVA problem. Only 71% who remembered the PVA formula could use it to solve for the annuity payment. Also, 95% of those who remembered the PVP formula could use it to correctly answer the PVP problem.

The heart of this study is to assess student perceptions of the derivations of the annuity formulas. Part A of Table 1 reports the number (and percentage) of students who correctly remembered each of the three forms. Part B of Table 1, reflecting the information presented in Part A, reports the number (and percentage) of students answering the problems correctly using each of the three forms. The majority of students favored Form C, followed by Form A and then Form B in all three types of annuity problems. The proportion of students who used Form B in computing the PVA (15%) is significantly less than those who used either Form A (38%) or Form C (48%). No significant difference is found between the proportion of students who used Form A and those who used Form C. However, when used to compute the FVA or the payment for the PVA, Form C is used significantly more than either of the other two forms.

Table 1C shows that while students who used Form A were the least able to calculate the correct answers for the PVA problem, no significant difference exists in the correct use of each form. Although a significant proportion of students used Form C in the FVA problem, the proportion of students correctly using either Form A or Form B was significantly greater than those using Form C. In other words, more students used Form C, but a larger proportion of students used Forms A and B correctly. The form that provided the best performance in calculating the payment for the PVA problem was Form C, followed closely by Form B. This seems reasonable, since using Form C or Form B involves less manipulation of the PVA formula to isolate the annuity payment (PMT).

Table 1 **Quiz Results**				
A1.	PVA	FVA	PMT for PVA	PVP
Total Attempts	86	86	86	86
Total correct formulas / total attempts (%)[a]:	61/86 (71%)***	62/86 (72%)***	59/86 (69%)***	77/86 (90%)***
A2.	PVA	FVA	PMT for PVA	PVP
Total correct forms / total correct formulas (%):				
Form A	23/61 (38%)***	16/62 (26%)***	17/59 (29%)***	-
Form B	9/61 (15%)***	10/62 (16%)***	10/59 (17%)***	-
Form C	29/61 (48%)***	36/62 (58%)***	32/59 (54%)***	-
Pair-wise Comparisons[b]:				
Form A vs. Form B	38% >15%***	26% > 16%***	29% > 17%***	-
Form A vs. Form C	38% < 48%***	26% < 58%***	29% < 54%***	-
Form B vs. Form C	15% < 48%***	16% < 58%***	17% < 54%***	-
B1.	PVA	FVA	PMT for PVA	PVP
Total correct problems / total formulas (%)[a]:	52/61 (85%)***	56/62 (90%)***	42/59 (71%)***	73/77 (95%)***
B2.	PVA	FVA	PMT for PVA	PVP
Total correct problems / total correct form (%):				
Form A	18/23 (78%)***	16/16 (100%)**	7/17 (41%)***	-
Form B	8/9 (89%)***	10/10 (100%)**	8/10 (80%)***	-
Form C	26/29 (90%)***	30/36 (83%)**	27/32 (84%)***	-
Pair-wise Comparisons[b]:				
Form A vs. Form B	78% < 89%***	100% ≠ 100%***	41% < 80%***	-
Form A vs. Form C	78% < 90%***	100% >83%***	41% < 84%***	-
Form B vs. Form C	89% < 90%***	100% > 83%***	80% < 84%***	-
***, **, * indicate significance at the 1, 5, and 10 percent, respectively				
[a] Chi-Square test is used				
[b] Marascuillo procedure is used				

Table 2 reports students' recall of the formulas and their perceptions of how the formulas were presented in the classroom. Students were asked if they benefited from the annuity formula derivation lecture and were also asked the write down the formulas for PVA, FVA, and PVP. Seventy-seven students were surveyed in total. A large number of students (24 out of 77, or 31%) did not state their perceptions while others gave either an unclear or neutral response. The vast majority had a favorable impression of the classroom presentation; 47 students (61%) stated that they were positively affected from having the derivation of the formulas presented to them in class. Only 6 students (8%) stated that they were negatively affected.

Table 2 **Survey Results**				
A.	Positive	Negative	N/A*	Total
Total Response	47/77 (61%)	6/77 (8%)	24/77 (31%)	77
Pair-wise Comparisons[b]:				
Positive vs. Negative	61% > 8%***			
Positive vs. N/A	61% > 31%***			
Negative vs. N/A	8% < 31%***			
B1.	Positive	Negative	N/A*	Total
PVA: Total Formulas Recalled	30/47 (64%)	1/6 (17%)	15/24 (63%)	46/77(60%)
Pair-wise Comparisons[b]:				
Positive vs. Negative	64% > 17%***			
Positive vs. N/A	64% > 63%***			
Negative vs. N/A	17% < 63%***			
B2.	Positive	Negative	N/A*	Total
PVA: (Total correct formulas / total response)				
Form A	8/30 (27%)	0/1 (0%)	2/15 (13%)	10/46 (22%)
Form B	6/30 (20%)	0/1 (0%)	0/15 (0%)	6/46 (13%)
Form C	16/30 (53%)	1/1 (100%)	13/15 (87%)	30/46 (65%)
Pair-wise Comparisons[b]:				
Form A vs. Form B	27% > 20%***	0% ≠ 0%***	13% > 0%***	22% > 13%***
Form A vs. Form C	27% < 53%***	0% < 100%***	13% < 87%***	22% < 65%***
Form B vs. Form C	20% < 53%***	0% < 100%***	0% < 87%***	13% < 65%***
C1.	Positive	Negative	N/A*	Total
FVA: Total Formulas Recalled	26/47 (55%)	1/6 (17%)	11/24 (46%)	38/77 (49%)
Pair-wise Comparisons[b]:				
Positive vs. Negative	55% > 17%***			
Positive vs. N/A	55% > 46%***			
Negative vs. N/A	17% < 46%***			
C2.	Positive	Negative	N/A*	Total
FVA: (Total correct formulas / total response)				
Form A	8/26 (31%)	0/1 (0%)	1/11 (9%)	9/39 (23%)
Form B	6/26 (23%)	0/1 (0%)	1/11 (9%)	7/39 (18%)
Form C	12/26 (46%)	1/1 (100%)	9/11 (82%)	23/39 (59%)
Pair-wise Comparisons[b]:				
Form A vs. Form B	31% > 23%***	0% ≠ 0%***	9% ≠ 9%***	23% > 18%***
Form A vs. Form C	31% < 46%***	0% < 100%***	9% < 82%***	23% < 59%***
Form B vs. Form C	23% < 46%***	0% < 100%***	9% < 82%***	18% < 59%***
D.	Positive	Negative	N/A*	Total
PVP: Total Formulas Recalled	39/47 (83%)	1/6 (17%)	4/24 (17%)	44/77 (57%)
Pair-wise Comparisons[b]:				
Positive vs. Negative	83% > 17%***			
Positive vs. N/A	83% > 17%***			
Negative vs. N/A	17% ≠ 17%***			
* Includes no response to sentiment or those responses that were either unclear or neutral in regards to sentiment				
***, **, * indicate significance at the 1, 5, and 10 percent, respectively				
[a] Chi-Square test is used				
[b] Marascuillo procedure is used				

Table 2 also shows that those students who were positively affected by the derivation presentation were better able to recall formulas than those students who were negatively affected.

Sixty-four percent of the positively affected students were able to recall the PVA formula (in various forms) compared to only 17% of those who were negatively affected. The pattern holds for the FVA (PVP) formula with 55% (83%) of the positively affected students able to recall the FVA (PVP) formula compared to only 17% (17%) for those negatively affected. Those who liked the presentation and were able to recall the PVA formula preferred Form C (Table 2.B2), but no significant difference exists between the forms for the FVA formula (Table 2.C2).

CONCLUSION AND IMPLICATIONS

This paper attempts to measure student response to teaching the derivation of annuity formulas in the classroom by observing formula usage as well as surveying student sentiment. Overall, presenting annuity formula derivations seems to benefit students. Student usage of all three forms indicates that some benefit is associated with presenting all three forms although, depending on the type of problem presented, some forms are used more than others and the usage of some forms results in better performance. Also, those students who were positively affected by the derivation presentation are better able to correctly recall formulas than those negatively affected. This result may be interpreted in a variety of ways. On the one hand, since
those who were not able to correctly recall the formulas in any form stated they were negatively affected, an assumption may be drawn that only poorer performing students dislike the derivation presentation. In other words, only the "bad" students do not appreciate "good" teaching. On the other hand, one would expect that those who were negatively affected would not be able to recall the formulas.

Other factors that may affect an instructor's decision to include the derivation lecture include the amount of class time devoted to deriving annuity formulas as well as the benefits and costs of using the various forms throughout the semester. After presenting the formula derivations, an instructor may illustrate how to calculate annuity problems with either all three forms or with only one form. For instance, an instructor may show how to calculate the present value of a bond's coupons (which form an annuity) with each of the three forms or with only Form C as is commonly the case. Of course, using all three forms is time consuming, and may also confuse a few students. With the increasing technology of online teaching platforms, an alternative approach would be to present the derivation of the formulas in video or PowerPoint format as supplemental material via the internet.

In summary, deriving the FVA and PVA annuity formulas results in three common forms of each of the formulas. Results from a quiz requiring the use of one of the three forms for each formula reveals that some forms are used more than others. The form that provides better performance depends on the type of problem presented. In addition, the majority of students indicated a positive sentiment toward presenting the derivation of annuity formulas in the classroom while a small minority disliked the derivation presentation. The results presented in this paper allow the instructor some insight into student perceptions of how annuity formulas should be taught.

REFERENCES

Brealey, R., S. Myers, and F. Allen (2014). *Principles of Corporate Finance (Eleventh Edition).* McGraw Hill Irwin.

Brigham, E.F. and M.C. Ehrhardt (2014). *Financial Management: Theory and Practice (Fourteenth Edition).* South-Western.

Newcombe, N.S., N.Ambady, J. Eccles, L. Gomez, D. Klahr, M. Linn, K. Miller, and K. Mix (2009). Psychology's Role in Mathematics and Science Education. *American Psychologist*, 64(6), 538-550.

Ross, S., R. Westerfield, and J. Jaffe (2013). *Corporate Finance (Tenth Edition).* McGraw-Hill Irwin.

Sherritt, L.W. (1944). Simplifying the Presentation of Compound Interest Formulas *Accounting Review*, 19(3), 310-314.

Skinner, F. S. (1994). An Integrative Approach to Teaching Valuation Equations. *Financial Practice and Education*, 4(1), 87-90.

Watson, J. D. (1936). Explaining Annuity Formulas. *Accounting Review*, 11(4), 388-389.

Welch, Ivo. (2009). *Corporate Finance: An Introduction.* Prentice Hall.

HOW DO INSTRUCTOR'S ATTENDANCE POLICIES INFLUENCE STUDENT ACHIEVEMENT IN PRINCIPLES OF MICROECONOMICS?

Todd Broker, Murray State University
Martin Milkman, Murray State University
Victor Raj, Murray State University

ABSTRACT

This paper examines the question of how an instructor's attendance policy influences student performance in Principles of Microeconomics. This study asked students in several different microeconomics classes at a medium sized regional university what sort of attendance policy they were subject to: was there a grade incentive for coming to class (i.e. bonus points), was there a grade punishment for not coming to class (i.e. deduction of points, missed assignments, etc.), was there some combination of the previous two or was there simply no attendance policy. While there are a few papers showing a positive correlation between required attendance and course performance, this paper seeks to understand more about the impact from the type of attendance policy employed. Data is collected from a student survey and from the university's registrar. The main empirical evidence is gathered from a two-stage regression analysis with student absenteeism as the dependent variable in the first equation and a student's final grade (using a 4.0 scale) as the dependent variable in the second equation. We find that, everything else equal, students seem more motivated to come to class when they expect a positive reward and they are more likely to miss class if they expect a negative punishment. Also, student attendance is a small, but significant determinant of a student's course performance after controlling for other relevant factors.

INTRODUCTION

The relationship between attendance and performance is an important issue in any classroom setting, including economics courses. Several studies have shown how attendance can influence student performance in the economics classroom (Chen and Lin, 2008; Stanca, 2006). Several other studies have shown evidence of how attendance policies can affect student attendance (Romer, 1993; Durden and Ellis; 1995; Marburger, 2001 & 2006; Dobkin et al., 2009). But, relatively few studies have focused attention on the question of whether the type of attendance policy implementation influences attendance and measured performance in economics (Self, 2012). This is the question we are asking in this paper.

We see from our data (Figure 1) that student absenteeism varies among the different attendance implementations. This graph shows that on the surface (i.e. without controlling for anything) students seemed less likely to miss class when faced with a positively incentivized

attendance policy compared to any other approach. Furthermore, we see preliminary evidence of an inverse relationship between absences and student performance (Figure 2).

Figure 1: Average Absence Percentage by Policy Type

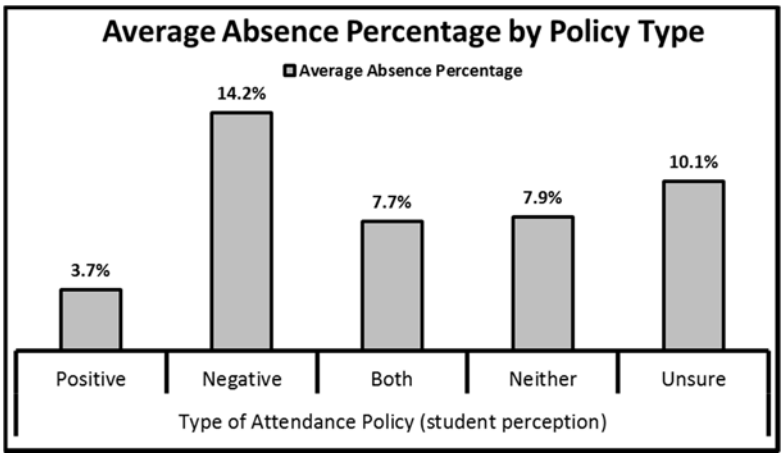
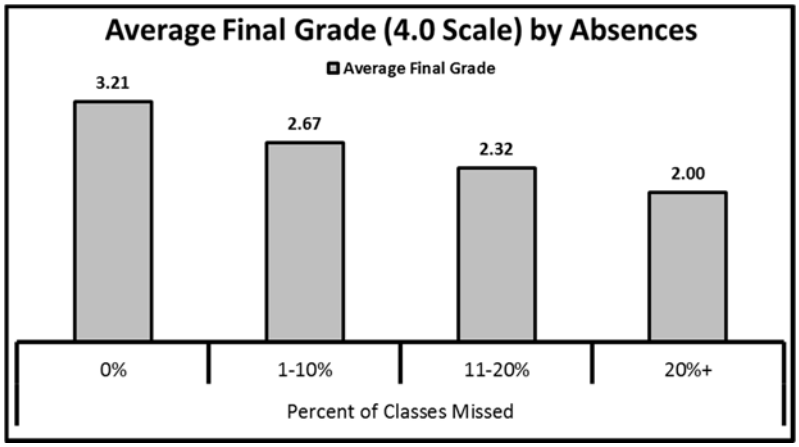


Figure 2: Average Final Grade by Absences



To investigate this preliminary evidence further, we used a two-stage regression approach. We first looked at whether using different types of attendance policies can have different effects on student absenteeism. Secondly, we looked at whether student attendance has any significant effect on student performance as measured by a student's overall course grade.

From here, the paper is divided into several parts. After this introduction, a brief review of the previous literature is presented. Then the data for this project is discussed and finally the results of the paper are offered with some modest conclusions.

REVIEW OF THE LITERATURE

The quantitative study of attendance and student performance in economics courses has a relatively short history, but also a relatively clear one: several studies have repeatedly found solid evidence that class attendance and performance share a strong, positive relationship.

Park and Kerr, in the spring of 1990, used a multinomial logit approach to study grades in different sections of a money and banking course. Using 97 observations and estimating only one equation, the authors find that absenteeism is far from the most important determinant, but does have a negative effect on grades in the class (Park and Kerr, 1990). This paper provides one of the first of several quantitative confirmations that student absenteeism negatively impacts student achievement in economics courses.

A second study by Romer looks at several intermediate macroeconomics classes. He first estimates an equation to measure absenteeism and then uses this equation to measure grades. He finds that generally, absenteeism negatively affects grades (Romer, 1993). Durden and Ellis also find that absenteeism significantly impacts student performance in principles of economics classes. They find a cumulative effect; "The evidence suggests that the effect is nonlinear, becoming important after a student has missed four classes during the semester" (Durden and Ellis, pg. 345, 1995).

Marburger, in 2001, allowed for a direct quantitative connection to be made between student attendance and learning. He "estimates a qualitative choice model in which the likelihood of responding incorrectly to a multiple choice question was related to whether the student was absent during the corresponding class period" when that particular material was covered (Marburger, pg. 100, 2001). He finds that absenteeism is a significant determinant of incorrect responses on multiple questions in a Principles of Microeconomics class. Marburger then extends this work in 2006 by doing a second paper where he analyzes two semesters of students. For one set of students their attendance was taken in class but it never affected their grade. For the other set of students, attendance did affect their grades. Marburger finds that the policy significantly reduced absenteeism. Then, following the same method of his previous research, he once again finds that being absent is a significant determinant of incorrect responses on exams (Marburger, 2006). This line of work creates an important first step towards making the connection between what an economics instructor can do to improve student attendance, and thus student learning.

Dobkin et al. looked at how enforcing a mandatory attendance policy on struggling students (determined by their midterm grade) for the second half of the course would affect attendance and performance. They find that student absenteeism falls significantly among the students exposed to the attendance policy. Also, among those same students, not only do they do better on their final exam in economics, but there is also not an adverse "crowding-out" effect, negatively harming their grades in other courses taken concurrently (Dobkin et al., 2009).

More recent research by Self, 2012, has investigated the impact of using different attendance policy strategies on student attendance. Her study focused on students in a Principles of Macroeconomics class. She found "strong support" that enforcing an attendance policy in class encourages students to come to class more often. She also found that students seemed more influenced to come to class if there was a negative punishment for missing class as opposed to positive reward for coming to class. This result is striking because it seems to contradict some of our main conclusions about the effect of positive versus negative incentives in attendance policy implantation. However, we believe our two studies can be reconciled in that the attendance policies she studied carried heavier punishments that the policies we studied, which could explain the difference in our results.

Either way, the past evidence seems overwhelmingly in support of the idea that students need to come to class to improve their learning of economics. Although there are some past studies that attempt to connect instructor enforced attendance policies with student attendance and subsequently performance, very few compare different types of attendance policies directly. We hope this study is another step in that direction.

DATA

This paper took advantage of a natural experiment in the Principles of Microeconomics sections at a comprehensive, medium-sized, public university in the south. Most of our data is self-reported from a voluntary survey with some supporting data provided by the Registrar.

In the spring of 2011, there were six sections of Principles of Microeconomics being taught by five different instructors [1]. Two of these sections used no official attendance policy, two other sections used attendance policies that effectively punished student's that didn't attend (through missing in-class quizzes and homework assignments) and the other two sections used an attendance policy that effectively benefited the students' final grade if they had perfect attendance (1.5 percentage points were added to their final overall percentage grade for perfect attendance).[2]

During that semester, 189 students enrolled in all six sections combined. The students were offered a voluntary survey near the end of the semester[3]. The average attendance rate on the dates the survey was offered was nearly 85%. This meant about 160 students were given the opportunity to take the voluntary survey. Of this group, about 93% took the survey, which gave us a sample size of 148 respondents (which amounted to an overall response rate of about 78% = 148/189). Part of the difference in these numbers is due to the fact that of the 189 registered students at the beginning of the semester, some of them dropped or withdrew during the course of the semester. In addition to information about each student's extracurricular life from the survey, we received most of the academic information for each student from the registrar. Tables 1-3 offer short explanations and then descriptive statistics of the variables used in our empirical work.

Table 1. List of Independent Variables
Dependent Variable = Absenteeism
(1st Equation)

Variable	Description
GPA	The students' cumulative grade point average [4].
Commute Time to Campus	The average number of minutes it takes for the student to get to campus from their residence.
Hours of Work	The average number of hours the student worked at a job in a given week during the semester.
Time Spent with Student Organizations	The average number of hours the student participated in activities of a student organization in a given week during the semester.
Tuition Payment Structure	A dummy variable equal to 1 if the student indicated that they paid at least part of their tuition with their own money or money they would have to pay back (student loan).
First or Last Class of Day	A dummy variable equal to 1 if the student indicated that their Principles of Microeconomics class was either their first or last class on the days it met.

Positive Attendance Policy	A dummy variable equal to 1 if the student indicated that they thought their class attendance policy included positive benefits for coming to class.
Negative Attendance Policy	A dummy variable equal to 1 if the student indicated that they thought their class attendance policy included negative consequences for missing class.
Neither Positive or Negative Attendance Policy	A dummy variable equal to 1 if the student indicated that they thought their class attendance policy included neither positive nor negative effects for attending or not attending class.

Table 2. List of Independent Variables
Dependent Variable = Final Course Grade
(2nd Equation)

Variable	Description
GPA	The student's cumulative grade point average[5].
Credit Hours	The number of college credit hours the student had earned.
Interest Level in Economics	The student's rating, on a 5-point scale, of how interested they were in economics (5 highest).
Predicted Percent of Total Absences	Predicted absenteeism rate of the student missed; predicted dependent variable from first equation (transformed).

Table 3. Descriptive Statistics of All Variables

	All Students	
Variable	**Mean**	**S.D.**
Student Rate of Absenteeism in Economics (transformed: square root)[6]	**2.18**	**1.64**
GPA	**2.93**	**0.64**
Commute Time to Campus	**14.17**	**20.44**
Hours of Work	**13.46**	**13.80**
Tuition Payment Structure (pct. affirmative)	**0.45**	**.50**
Time Spent with Student Organizations	**6.56**	**9.55**
Interest Level in Economics	**2.95**	**1.10**
First or Last Class of Day (pct. affirmative)	**0.78**	**0.41**
Credit Hours	**75.88**	**35.51**
Student Final Course Grade (4.0 scale)	**2.69**	**0.99**
Predicted Student Rate of Absenteeism in Economics [7]	**1.83**	**0.92**
Positive Attendance Policy (pct. affirmative) [8]	**0.59**	**0.49**
Negative Attendance Policy (pct. affirmative)	**0.39**	**0.49**
Neither Positive or Negative Attendance Policy (pct. affirmative)	**0.20**	**0.40**
Number of Students	**148**	

METHODS

This paper sought to understand the empirical relationship between different types of attendance policies employed in an economic classroom and their effect on absenteeism, while controlling for other relevant factors. The second goal was to understand the empirical relationship between a students' absenteeism and their performance in the course, while controlling for each student's personal effort, intelligence, attitude, etc.

We utilized OLS regression techniques to estimate our parameters in the first equation. This seemed consistent with the prior literature (Dobkin et al., 2009) and with our own thoughts about the nature of individual student absenteeism, which served as our dependent variable.[9]

It was a critical concern for us in deciding how to structure the independent variables of interest (e.g. the types of attendance policies). We first considered assigning each student a dummy variable indicating what type of policy they were exposed to, based on the instructor's syllabus[10]. This however would have given us a dataset that was essentially aggregated at the class level. Since we only had six total classes, we essentially would have only had six unique data points to estimate our parameters[11]. Thus, we decided to use the self-reported data from the student surveys at the individual level. This would mean we would be assessing what each student's perception of their attendance policy was, as opposed to what it actually was according to the syllabus.

This, of course, created the possibility that the students could have misinterpreted the attendance policy in their class (Our data indicates that 67% of students in our study correctly reported their classes' attendance policy according to the syllabus). However, this seemed reasonable for our purposes given that we were empirically studying student responses to attendance policies, which would inherently be considering student perception. It could be true that it doesn't really matter what any actual attendance policy is, but rather the important factor could be in effectively communicating the policy to students so that they understand. Using the "student perception" approach in this study actually may be a better way to investigate this effect.

To measure each student's perception of their class attendance policy, they were asked what they thought the attendance policy was in their class. They had the option of responding that it provided a positive benefit, a negative consequence, both positive and negative effects, neither a positive nor a negative effect or they could have indicated that they didn't know what the attendance policy was. This information was used to construct our attendance policy dummy variable in the first equation where the dependent variable was absenteeism.

Given the fact that grades are reported to the university as discrete variables (A, B, C, D, E), OLS would not be sufficient to estimate student performance in our second equation. Therefore, to estimate the major determinants of a student's course grade we employed an ordered logit function where our dependent variable followed the 4.0 scale (A = 4, B = 3, C = 2, D = 1, E = 0).

Table 4. OLS Regression Statistics: Dependent Variable = Absenteeism (1st Equation)[14]			
Multiple R	0.612		
R Square	0.374		
Adjusted R Square	0.334		
Standard Error	1.337		
Observations	148		
	Coefficients	***Standard Error***	***t Stat*** [15]
Intercept	4.384	0.655	6.695**
GPA	-0.907	0.175	-5.167**
Commute Time to Campus	0.004	0.006	0.771
Hours of Work	0.014	0.009	1.511
Time Spent with Student Organizations	0.039	0.013	3.018**
Tuition Payment Structure	-0.657	0.247	-2.663**
First or Last Class of Day	0.699	0.280	2.497**
Positive Attendance Policy	-0.948	0.308	-3.078**
Negative Attendance Policy	0.763	0.252	3.029**
Neither Attendance Policy	-0.204	0.372	-0.548

RESULTS

This paper attempts to study the effects of various attendance policies empirically. The 1st equation of results, where the dependent variable was absenteeism, showed that our overall model performance was significant (as measured by the F-Statistic and R^2) and most of the other outcomes seemed generally consistent with previous research (see Table 4). Consistent with nearly all prior literature, a student's GPA[12] was significant and negatively correlated to absenteeism, meaning that students with higher overall GPA's were more likely to attend class. As already observed in previous studies, a student's GPA is probably a proxy not just for their intelligence, but also for a level of effort as well[13].

In addition, we see that students that spend more time (weekly hours) in campus organizational activities are significantly more likely to miss class. Interestingly, if a student indicated that that they were paying for tuition themselves, either through direct payments or indirectly through loans, they were significantly more likely to attend class. Not surprisingly, if a student indicated that their microeconomics class was their first or last class of the day for them personally, they were significantly more likely to miss class.

In terms of policy analysis, we see that both "positive" and "negative" attendance policies were significant predictors of absenteeism. However, only the positive attendance policy variable showed a negative relationship to absenteeism. In other words, students who thought there was a positive incentive for coming to class were more likely to go to their microeconomics class,

everything else equal. But those students that thought that there was a negative punishment for missing class were significantly more likely to miss class.

This outcome could be a result of the "positive" attendance policy offering an exogenous reward (i.e. bonus points) for coming to class while our sample of "negative" attendance policies mainly offered an endogenous punishment (i.e. missed quizzes/homework grades).

The results from the 2nd equation, where the dependent variable was a student's final course grade, returned a significant model (see Table 5). Our main variable of interest, student absenteeism, showed a significant negative relationship to a student's final performance in their Principles of Microeconomics class. This means that, even after controlling for student maturity (Credit Hours), student interest (Interest Level in Economics) and intelligence & effort, (GPA) students who came to class more often were significantly more likely to perform better (i.e. get a higher grade).

However, the coefficient on our variable of interest, absenteeism, was relatively small as were its marginal effects. This may mean that student absenteeism, though significant, may not have a meaningful impact on student performance until relatively large numbers of absences are accumulated (see Table 6) [16]. However, this result does not negate the fact that missing more class does generally harm a student's overall grade, everything else equal.

Table 5.
Logit Regression Statistics:
Dependent Variable = Final Course Grade
(2nd Equation)

Number of observations	148		
Log likelihood function	-144.136		
Restricted log likelihood	-201.314		
Chi squared	114.356		
Degrees of freedom	3		
Prob [ChiSqd > value] =	0.0000000		
	Coefficients	***Standard Error***	***B / St.Er.***[17]
GPA	2.120	0.190	11.133**
Credit Hours	0.009	0.005	1.901
Interest Level in Economics	0.268	0.141	1.906
Predicted Student Rate of Absenteeism in Economics	-0.194	0.050	-3.896**

Table 6. Summary of Marginal Effects for Ordered Probability Model (logit)					
	Y = 00 (Grade = E)	**Y = 01 (Grade = D)**	**Y = 02 (Grade = C)**	**Y = 03 (Grade = B)**	**Y = 04 (Grade = A)**
GPA	*-0.0022*	*-0.0715*	*-0.3895*	*0.2543*	*0.2089*
Credit Hours	*0.0000*	*-0.0003*	*-0.0016*	*0.0010*	*0.0009*
Interest Level in Economics	*-0.0003*	*-0.0090*	*-0.0493*	*0.0322*	*0.0264*
Predicted Student Rate of Absenteeism in Economics	*0.0002*	*0.0066*	*0.0357*	*-0.0233*	*-0.0192*

As expected, a student's overall GPA was significant and positively related to their final grade in economics. This suggests that students who have performed well in their other classes during college are likely to also perform well in their Principles of Microeconomics class, everything else equal.

Also, students with higher class standing (i.e. more college credit) and those that were more interested in economics were more likely to go to class as indicated with positive coefficients. But these effects were only marginally significant at the 90% level.

CONCLUSIONS

We set out to study empirically if student attendance was a significant contributor to student performance in Principles of Microeconomics classes. We also wanted to know if using an attendance policy in the classroom would help encourage attendance, and if so, what type of attendance policy seemed to have the largest impact in reducing student absenteeism.

Anecdotally we believed there was some extra motivation for students when they believed they could "earn" bonus points for their final grade in economics simply by coming to class. Despite recognition of possible empirical deficiencies, our results seem to suggest that students may in fact be more motivated to attend class when they think there is some sort of tangible reward for attending (i.e. bonus points). This seems striking when you consider that the bonus points that were offered in this experiment were really inconsequential (most teachers round their grades at the margins anyway).

Our result that positive attendance rewards may be a better motivator to reduce student absenteeism may not completely contradict previous studies that found that a negative punishment was a stronger motivating factor in deterring absenteeism. The punishment in those studies was seemingly quite severe: 5 points deducted from the students' overall points earned for every week of class missed not to mention missed quiz and homework grades (Self, 2012). By contrast, the negative attendance policy we studied only factored in missed quiz and homework grades with no explicit grade punishment. It could be possible that any reward or punishment that is tangibly explicit and large could be effective in encouraging class attendance.

ENDNOTES

1. Using five different instructors may cause unobserved heterogeneity in teaching styles / grading procedures that may bias our results. We attempted to control for these potential unobserved effects by controlling for the student interest level in the course. This variable showed statistically different values between courses/instructors suggesting that it might be serving as a proxy for some unobserved differences between the instructors. Furthermore, we do not believe that varying grading procedures have unduly influenced our sample since none of the different class averages (final grades) were statistically different from the overall sample average.
2. These classes were not random, per se, in the sense that students were allowed to self-select which classes they wanted to be in and the individual instructors were all free to select their own attendance policies at will. However, it would have been difficult for students to know which classes would offer which type of attendance policy before selecting one of these sections. For this reason, we do not feel that the data suffers unduly from self-selection bias.
3. The date the survey was offered in each section was coordinated with each instructor so as not to be inadvertently right before or after an exam, which may have biased our data collection inadvertently.
4. This is the student's cumulative GPA at the end of the semester they were taking their Principles of Microeconomics course. This isn't perfect, but it was the only cumulative data we were able to obtain from the Registrar.
5. Ibid.
6. In our final model we used to estimate our parameters, we transformed our dependent variable, student rate of absenteeism, by taking the square root. We did this to correct for skewness in the variable which allowed us to meet the OLS normality assumptions in our 1st equation. We ran the model both ways (with and without the transformed dependent variable) and recorded no significant differences in the results. The statistics in this table represent the variable after it was transformed.
7. The dependent variable in the 1st equation was transformed by taking the square root to correct for skewness in the data. Thus, our predicted values for absenteeism represented square roots. The numbers in this table were squared before being used in the 2nd equation, which is what the values in this table represent.
8. The attendance policy variables (self-reported by the students) add up to more than 100% because students were allowed to answer that they thought their policy had BOTH positive and negative incentives, which was represented in our equations as an affirmative response to both the "positive" and "negative" dummy variable.
9. The absenteeism variable used in our analysis was the self-reported number of absences each student reported missing on the survey divided by the number of classes that section met during the semester. To ensure accuracy, this was crossed checked and found to be highly correlated to attendance data collected by individual instructors. The correlation coefficient between the two datasets was 0.90.
10. We used this approach and did our empirical work using basically the same techniques and came to similar, but somewhat different results. Being exposed to a positive attendance policy still significantly encouraged attendance and attendance was a minor, but still significant, predictor and class performance. One difference in this approach was that the sign on the coefficient for the negative attendance policy was negative, like on the positive attendance policy dummy variable, indicating that it too reduced absenteeism. This sign was reversed when we used individual level data as opposed to class aggregated data.
11. Thanks to Caleb Stroup of Vanderbilt University for his helpful comments in this area.
12. The GPA was the cumulative GPA for the student after the semester was over.
13. Park and Kerr, pg. 105, 1990.
14. This first OLS regression model returned a significant F statistic of 9.175.
15. *statistically significant at the 95% level
    **statistically significant at the 99% level
16. As suggested by Durden and Ellis, pg. 345, 1995.
17. *statistically significant at the 95% level
    **statistically significant at the 99% level

REFERENCES

Chen Jennjou and Tsui-Fang Lin. 2008. Class Attendance and Exam Performance: A Randomized Experiment. Journal of Economic Education 39 n3.

Dobkin, Carlos, Ricard Gil, and Justin Marion. 2010. Skipping class in college and exam performance: Evidence from a regression discontinuity classroom experiment. Economics of Education Review 29 n4: 566-575.

Durden, Gary C., and Larry V. Ellis. 1995. The Effects of Attendance on Student Learning in Principles of Economics. American Economic Association 85 n2: 343-346.

Kerr, Peter M., and Kang H. Park. 1990. Determinants of Academic Performance: A multinomial Logit Approach. Journal of Economic Education 21 n2: 101-111.

Marbuger, Daniel R. 2001. Absenteeism and Undergraduate Exam Performance. Journal of Economic Education 32 n2: 99-109.

Marbuger, Daniel R. 2006. Does Mandatory Attendance Improve Student Performance? Journal of Economic Education 37 n2: 148-155.

Romer, David. 1993. Do Students Go to Class? Should They? Journal of Economic Perspectives 7 n3: 167-174.

Self, Sharmistha. 2012. Studying Absenteeism in Principles of Macroeconomics: Do Attendance Policies Make a Difference? Journal of Economic Education 37 n3:251-266

Stanca, Luca. 2006. The effects of attendance on academic performance: Panel data evidence for introductory microeconomics. Journal of Economic Education 37 n3:251-266

EXPLORING DATA MINING AND GAMIFICATION AS TOOLS FOR POVERTY ANALYSIS AND POLICY FORMULATION: A METHODOLOGICAL FRAMEWORK

Remedios de Dios Bulos, De La Salle University
Neriza M. Delfino, De La Salle University
John Paolo R. Rivera, De La Salle University

ABSTRACT

Economies are deemed to have automated systems of collecting their respective poverty statistics because the availability of these databases are imperative, which have generated a need for new techniques that will transform the data into useful information that can assist policymakers in formulating effective poverty-reduction policies. An alternative method is the use of data mining techniques – the extraction of hidden patterns of information from large databases, beyond regression, that will allow for the generation of a prediction on the direction and extent of change on the status of a sample. After determining these hidden patterns, the application of gamification can be utilized as a behavior change mechanism particularly for people who are predisposed to engaging in games on issues that are reality-based. The change in behavior gets manifested in their opinions about an issue or problem and subsequently takes positive actions towards a solution to a problem. Various problems besetting poverty stricken households can be solved through change in behavior once they realize the opportunities that can come their way by doing so. The change is much more effective if part of the motivation comes from within the individual and the immediate environment makes it possible for them to see the positive effect this change will bring.

INTRODUCTION

Poverty remains one of the rampant social problems any economy needs to address. With the Philippines' poverty line marked at earnings less than PHP 16,841.00 per individual annually. According to data from the National Statistical Coordination Board (NSCB), 26.5% of the population falls below the poverty line in 2009. Though this figure is a much lower than a comparative figure of 33.1% in 1991, the decline has been slow and uneven, much slower than neighboring countries who experienced broadly similar numbers in the 1980s, such as the People's Republic of China (PRC), Thailand, Indonesia (where poverty level lies at 8.5%) or Viet Nam (13.5%). The Philippines' incidence of poverty remains significantly high as compared to other countries for almost a decade now. The unevenness of the decline has been attributed to a large range of income brackets across regions and sectors. Rapid population growth has also contributed to this predicament (Rivera & See, 2012).

The government planned to eradicate poverty as stated in the Philippines Development Plan (PDP). The PDP for the next six years are an annual economic growth of 7% to 8% and the achievement of the Millennium Development Goals (MDGs). Under the MDGs, Philippines committed itself to halving extreme poverty from a 33.1% in 1991 to 16.6 % by 2015 (Rivera, Pizarro, Aliping & Reyes, 2012).

Poverty assessment and monitoring is one of the focus areas of the United Nations Development Program. According to the United Nations Development Programme [UNDP] (2012), fighting poverty and making progress towards the MDGs requires effective policies to reduce poverty and promote inclusive development. The design and implementation of policies for social inclusion require a good system of information to better understand the problem: what is poverty, what causes it, who does it affect, how does it evolve over time and what impact do development policies and programs have on poverty. Establishing a poverty monitoring and assessment system to answer these questions is fundamental to the design of effective poverty reduction policies. Such systems are also necessary to help governments and the development community to keep track of progress towards the MDGs (UNDP, 2012).

According to the National Anti-Poverty Commission [NAPC] (2012), fiscal constraints in the Philippines have compelled the government to implement targeted interventions directing public resources to the poor and marginalized groups. Identifying who and where the poor are is vital in efficiently and effectively implementing poverty-alleviation programs and projects. The Department of Social Welfare and Development (DSWD) is adopting the National Household Targeting System for Poverty Reduction (NHTS – PR) using the Proxy Means Test (PMT) in identifying their target beneficiaries for their flagship program, the *Pantawid Pamilyang Pilipino Program* (4Ps). The local government units (LGUs) have been encouraged to utilize the Community Based Monitoring System (CBMS) through SDC Resolution No. 3, Series of 2006 and Cabinet Secretary's Memorandum dated 11 March 2008 to target beneficiaries for their poverty programs and to evaluate these programs.

The NAPC (2012) has emphasized that the NHTS-PR is a data management system that identifies who and where the poor households are. It generates to the public a database of poor households as reference in identifying beneficiaries of social protection programs. Likewise, the system is also envisioned to reduce inclusion of inadvertent beneficiaries and exclusion of intended beneficiaries of social protection programs.

The NHTS-PR utilizes a "paper and pencil" approach in gathering data. It evaluates all households in all deprived areas and those pockets of poverty via house-to-house interviews. It collects information from the Household Assessment Form – a two-page questionnaire with 34 variables of interest. It applies the Proxy Means Test (PMT) – model, which according to NAPC (2012), estimates poverty level of households and ranking them based on provincial poverty thresholds. It is administered using standardized software developed to allow easy input of the household assessment information and homogenous processing of the information. This is to guarantee the quality of the information generated by the system.

A statistical formula computes the households' approximate income using proxy socio-economic variables that would predict household income and allow for objective ranking and classification of households into non-poor, survival poor, and food-poor. The selected variables considered good proxies of income are the following as per NAPC (2012): (1) household

consumption; (2) education of household members; (3) housing conditions; (4) access to basic services; (5) ownership of assets, tenure status; and (6) regional variables.

The CBMS Survey was developed to provide a good information base for policymakers in monitoring the effects of economic reforms to society's vulnerable groups. It addresses the inadequacy of vital yet disaggregated information for poverty analysis and design of appropriate interventions, for targeting of program beneficiaries and for program impact monitoring. CBMS' five objectives, as cited by NAPC (2012) are (1) to diagnose the extent of poverty at the local level; (2) formulate appropriate plans and programs to address problems; (3) provide the basis for rational allocation of resources; (4) identify eligible beneficiaries for targeted programs; and (5) monitor and assess the impact of programs and projects.

Note that the CBMS involves a census of households in a community. It is LGU-based and it fosters community participation by tapping existing LGU personnel/community members as monitors. It has a core set of indicators but the system can accommodate additional indicators. It establishes database at each geopolitical level (NAPC, 2012).

The data, which could be generated from the CBMS, contain the core set of indicators that covers the multidimensional nature of poverty. The core indicators are: (1) Proportion of children aged 0 to 5 years old who died to the sum of children aged 0 to 5 years old; (2) Women deaths due to pregnancy-related causes; (3) Malnutrition prevalence/proportion of children aged 0 to 5 years old who are malnourished to the total number of children 0 to 5 years old; (4) Proportion of households without access to safe water; (5) Proportion of households without access to sanitary toilet facilities; (6) Proportion of households who are squatting; (7) Proportion of households who are living in makeshift housing; (8) Proportion of households with members victimized by crimes; (9) Proportion of households with income less than the poverty threshold; (10) Proportion of households with income less than the food threshold; (11) Proportion of households who eat less than 3 meals a day; (12) Unemployment rate; (13) Elementary school participation rate; and (14) Secondary school participation rate. Other local government (LGU)-specific indicators relating to disabilities, natural calamities, migration, waste management, access to programs, electoral participation and community organization (NAPC, 2012).

In the Philippines, there are at least two separate automated systems of collecting and storing poverty data that are currently operational – NHTS-PR and CBMS. The availability and accessibility of these databases have generated a compelling need for new techniques and automated tools that can intelligently assist policy-makers in formulating poverty reduction policies. Parallel to the idea of Han, Kamber and Pei (2011), new and better ways to automatically analyze, classify, summarize, discover and characterize trends and flag anomalies in poverty data are urgently needed. Transforming the data into useful information and knowledge that will assist in the design of effective poverty reduction policies is imperative.

Data mining has been adopted by many organizations and has been utilized intensively and extensively as a decision-support tool. Yet to date, there has been a dearth of research work in using data mining as a tool in poverty assessment and analysis. Policy-makers often expend a lot of resources in the formulation of poverty reduction policies. However, due to lack of collective accountability, the best policies tend to fall through. Thus, it is important that decision-makers implement poverty reduction policies that are clear, understandable, realistic, consistent and enforceable. Policy-makers should utilize every possible means to convince the citizenry and other

concerned sectors of society of the benefits of the poverty reduction policy initiatives they have formulated. They should find the venue to effectively communicate policy information credibly. They should be able to regularly engage the citizenry and other sectors of society towards the successful adoption and implementation of the poverty reduction reforms.

The problem that will be tackled in this study is in determining the extent by which data mining can establish trends in the CBMS database or relationships among data gathered at the household level that would point to specific problem areas in the community. Once the problem areas are identified, information dissemination and behavior change efforts can be instituted by way of the appropriate game design. From this backdrop, this study can explore the following specific objectives.

- Identify problem areas in a specific low-income community using the existing database collected from households (i.e. CBMS) in this area through data mining;
- Develop game designs that would get various stakeholders involved, firstly, in being aware about the community problem(s) and its root cause and, secondly, what appropriate actions they are willing to take towards solving the problem.
- Conduct pre- and post evaluation and cost-benefit studies of the interventions conducted to determine its effectiveness.

ALTERNATIVE APPROACHES TO POVERTY ANALYSIS

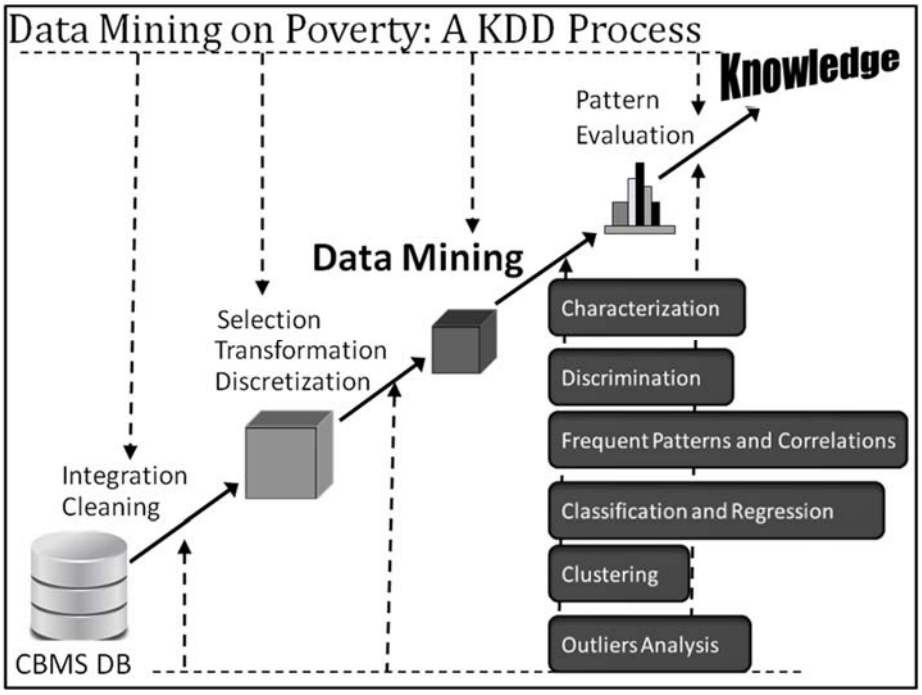
Data mining (DM), as seen from Figure 1, is the extraction of hidden patterns of information from large databases. It has been extensively used by various types of organizations: financial institutions, for credit scoring and fraud detection; marketers, for direct marketing and cross-selling or up-selling; retailers, for market segmentation and store layout; and manufacturers, for quality control and maintenance scheduling (Koh & Tan, 2005).

Basically, the goal of data mining is to generate either prediction on direction and extent of change on the status or condition of the target class or description of the subjects' current and past status. Predictive mining performs induction on the current data in order to make predictions. It involves using some variables or fields in the data set to classify, predict unknown values or estimate values of the variables of interest. On the other hand, descriptive mining characterizes properties of the data in a target data set. It involves finding patterns and relationships in the data that can be interpreted. For example, DM tools can spot the frequent occurrence of open dumpsites and pest infestation together (Han, Kamber & Pei, 2011; Syed, 2011).

There are a number of data mining functionalities, which are being used to specify the kinds of patterns to be found in data mining tasks. These include characterization and discrimination; the mining of frequent patterns, associations, and correlations; classification and regression; clustering analysis; and outlier analysis (Han, Kamber & Pei, 2011).

Data characterization, as defined by Han, Kamber and Pei (2011), is a summary of the general qualities of a target class of data. The data corresponding to the user-specified class are collected by a query. Meanwhile, data discrimination, as defined also by Han, Kamber and Pei (2011), is a comparison of the general features of the target class data objects against the general features of objects from one or multiple contrasting classes. A user can specify the target and contrasting classes and the corresponding data objects can be retrieved through database queries.

Figure 1
Data Mining on Poverty – The Process



Source: Adapted from Han, Kamber and Pei (2011)

Frequent patterns are patterns that occur habitually in data such as frequent item sets, sequential patterns, and frequent substructures. Mining of frequent patterns leads to the discovery of interesting associations within data (Han, Kamber & Pei, 2011).

On the one hand, classification is the process of finding a model that describes and distinguishes data classes. The models are derived based on the analysis of a set of training data. The model is used to predict the class label of objects for which the class label is unknown. Whereas, classification predicts categorical labels and regression models continuous-valued functions – regression is used to predict missing or unavailable numerical data values rather than discrete class labels. The term prediction refers to both numeric prediction and class label prediction. Regression analysis is a statistical methodology that is most often used for numeric prediction, although other methods also exist. Regression also comprises the identification of distribution trends based on the existing data (Han, Kamber & Pei, 2011).

On the other hand, clustering can be used to generate class labels for a group of data. The objects are clustered based on the principle of maximizing intra-class similarity and minimizing interclass similarity – clusters of objects are formed so that objects within a cluster have high similarity in comparison with each other, but are dissimilar to objects in other clusters. Each cluster formed can be viewed as a class of objects, from which rules can be derived. Clustering can also facilitate taxonomy formation – the organization of observations into a hierarchy of classes that group similar events together (Han, Kamber & Pei, 2011).

As accentuated by Han, Kamber and Pei (2011), a data may encompass objects that do not conform to the general behavior of the data – outliers. Various data mining methods discard

outliers as exclusions. However, the sporadic events can be more interesting than the more regularly transpiring ones. The analysis of outlier data is referred to as anomaly mining. Outliers may be detected using statistical tests that assume a distribution for the data, or using distance measures where objects that are remote from any other cluster are considered outliers. Rather than using statistical or distance measures, density-based methods may identify outliers in a local region, although they look normal from a global statistical distribution view.

Data mining is a powerful new technology with great potential Alexander (n.d.). With respect to poverty reduction, it can help policy-makers focus on the most important information in poverty databases. The application of data mining functionalities can enable the prediction of future trends and behaviors in poverty data, thereby allowing decision-makers to make proactive, knowledge-driven decisions towards poverty reduction and advancement of human development (Adeyemo & Kuye, 2006). Data mining tools can scour poverty databases for hidden patterns that poverty experts may miss because they lie outside their realm of regular expectations (Alexander, n.d.).

GAMIFICATION: TURNING POLICY INTO ACTION

According to Zicherman and Cunningham (2011), the idea of using game-thinking and game mechanics to solve problems and engage audiences isn't exactly new. The military has been using games and simulations for hundreds, if not thousands, of years, and the Unites States of America (USA) military has been a pioneer in the use of video games across branches. They also say that play and games are enshrined in our cultural record, emerging with civilizations, always intertwined; and that we are hardwired to play, with researchers increasingly discovering the complex relationships between our brains, neural systems, and game play. Gamification is the use of game design techniques, game thinking and game mechanics to enhance non-game contexts such as poverty. Typically it applies to non-game applications and processes, in order to encourage people to adopt them, or to influence how they are used. It works by making technology more engaging, by encouraging users to engage in desired behaviors, by showing a path to mastery and autonomy, by helping to solve problems and not being a distraction, and by taking advantage of humans' psychological predisposition to engage in gaming. The technique can encourage people to perform chores that they ordinarily consider boring, such as completing surveys, shopping, filling out tax forms, or reading web sites (Herger, 2012).

The gamification of poverty, an example of which is the game "Spent" by Nicholson (2011) as mentioned by Coren (n.d.) can show that "what games can accomplish is empathy." Acording to Mims (2011), it may be a vehicle to "allow the player to live a life of adverse poverty, through an immersive experience of tough decisions" (Lopez, 2011). It can bring all major players of poverty reduction together and through regular and continuous engagements, the participants may discover new paths out of poverty. It may encourage participants to give back to the less fortunate. It may assist activists in getting their message across. And in the end, it may perhaps motivate the less fortunate in our society to improve their lives and negotiate their way out of poverty.

It is the stage to engage and enable the participants to imagine the best-case scenario outcome for real-world problems such as poverty. It is a means by which people can be empowered to make the outcome a reality by giving them the means to achieve victory (McGonigal, 2010). Gamification as a complete concept is still in its infancy (Zicherman & Cunningham, 2011). It can

be challenging to think about how they can be applied to world-changing and world-saving applications such as poverty reduction. In terms of altering behavior through the use of gamification, an example cited in Coren (n.d.) and Nicholson (2012) is a "physical therapy visualization tool that allows the patient to see how the body is changing as he/she does each repetition can allow each patient to set a different goal that is meaningful. The therapist can help the patient set goals through constraints, and by exploring those constraints, the patient can understand how the physical therapy connects to the exercise goals. By giving the patient information and control over goals, the patient is much more likely to find the internal meaningful connections to be able to continue the therapy away from the therapist."

The benefits of gamification can also be seen in its impact on benefactors. One such game is the Community Impact Platform gamifying a user's experience where giving back becomes an engaging experience. The fund raising game competition "encouraged employees to create personal fund raisers that tap into their unique network. The companies they work for, in turn, challenge parameters, select non-profit targets, match donations and create customized pages to promote among their social media network, fostering a grassroots cause marketing for the company that also amplifies its social impact" (Scott, 2012).

METHODOLOGICAL FRAMEWORK

With the use of data mining techniques/procedures, identify trends and relationships between a community households' demographic make-up and the problems that beset the community. The succeeding procedures are as follows:

- Utilize such relationships to determine specific problem areas that can be addressed through community based action programs or policy reforms. If community based programs or policy reforms already exists, then there will be a need to enhance the effectiveness of delivery of service and dissemination of information on the benefits of availing public service through behavior-changing games.
- Survey on the extent of usage of Internet cafes and cell phones by a sample of households within the community, to determine if this mode of information dissemination as well as data collection is feasible within the community. The lack of usage of this form of technology may warrant the need for alternative forms of information dissemination that can be made available in barangay or health centers assisted by personnel adept with the use of the technology. The surveys will also have to cover the area of literacy due to the need for this to make the technology usable to the community, particularly with reference to gamification. Participants will then have to be divided into two groups: those who can do the game by themselves and those who will have to be assisted.
- A game is developed to address the issue affecting the community that would require behavior change. As in any gamification effort, the participant will have to be rewarded with some form of point system redeemable and exchanged into some basic need (like commodity) of the household. This would also entail the development of a reward system that would sustain the participants' interest in the game until the objective of the game is fulfilled.

- Post-game evaluation on behavioral change is conducted to determine effectiveness of the game in addressing the community's problems and the modifications needed to improve the game in fulfilling its purpose if it can be used in other communities.
- Consider the development of other games as may be deemed necessary to address other pressing issues within the sample community to sustain the success (if ever) and advantages put forth by previous gamification efforts.
- Introduce a data collection mechanism for data collectors (researchers), possibly also in the form of a game, to be able to extend the reach of data collection to as many households. This would involve the collection of quality data which is more accurate and which is updated constantly on a regular basis so that the community can easily move forth with programs attuned to immediate needs.
- A cost-benefit analysis is also to be conducted to compare the effectiveness of the intervention proposed and determine how the proposed intervention can complement existing services being delivered in the community.

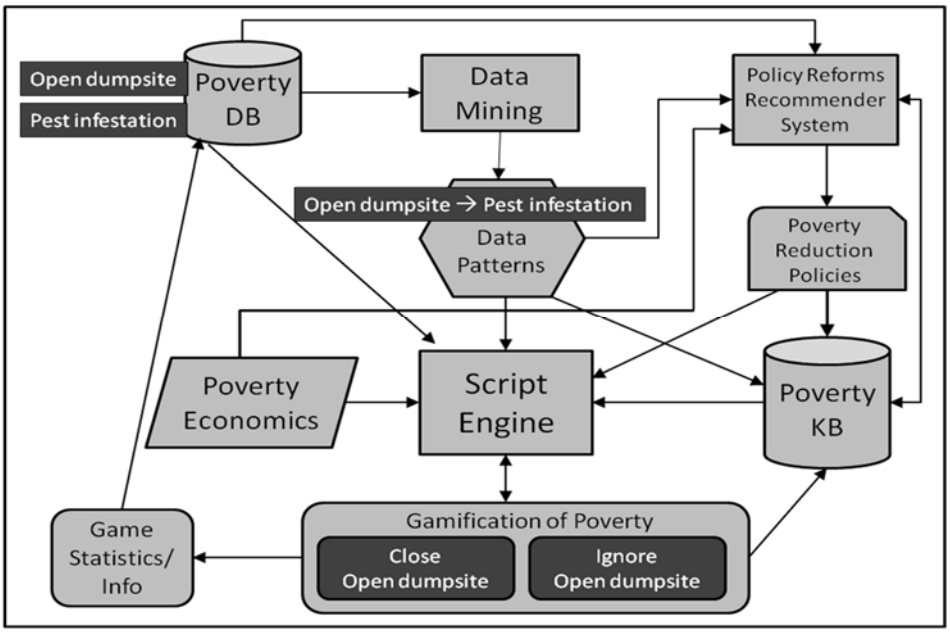
The prospects on behavior change that gamification brings can easily extend to areas that tend to plague informal dwellers and poverty stricken sectors that can assist in addressing health issues, fertility, sanitation, peace and order, employment, livelihood and subsequently, poverty.

The system of monitoring, assessing and data-update with the use of gamification and data-mining for policy research will be compared to the current system in terms of monetized as well non-easily monetized costs and benefits that accrue to the implementing unit. Some questions CBA try to answer include: (1) Does gamification and data-mining provide the government unit substantial benefits; (2) should implementation of the new technology be applied in other areas that serve the sector involved; (3) if there are other means to address the concerns of the community, this shall be compared with the proposed one. The CBA is also seen to be able to establish links between inputs and outputs, determine underlying assumptions on program implementation, as well as the identification of 'opportunity' cost with the non-implementation of the new technology.

The Proposed System and Data Mining

The proposed system architecture is shown in Figure 2. It consists of three main modules, namely, Data Mining, Policy Reforms Recommender System and Gamification of Poverty. The Data Mining module would involve data pre-processing, attribute selection, application of data mining algorithms, pattern evaluation and knowledge presentation. Researchers can utilize existing data mining system (i.e. WEKA or Rapid Miner) or develop their own algorithms as deemed appropriate to generate interesting data patterns. The application of various data mining functionalities can be investigated such as characterization and discrimination; the mining of frequent patterns, associations, and correlations; classification and regression; clustering analysis; and outlier analysis.

Figure 2
System Architecture



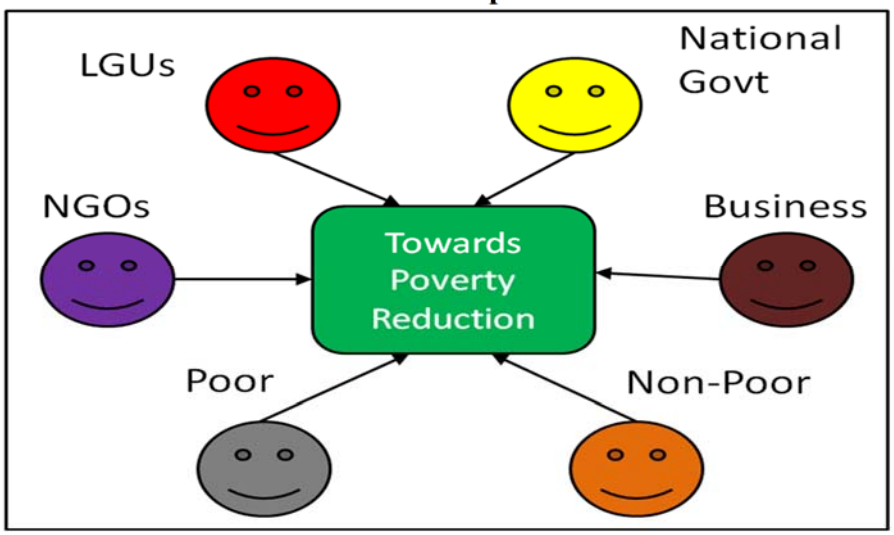
Policy Reforms Recommender System for Future Research

The Policy Reforms Recommender System will automatically and intelligently generate poverty reduction policies. It will attempt to identify different poverty reduction alternatives such as programs or spending priorities. It will use impact or cost-benefit analysis, and apply poverty economic principles to select and recommend which among those identified will be adopted. Machine Learning algorithms and Artificial Intelligence techniques will be investigated and applied in crafting poverty reduction reforms. These algorithms will intelligently analyze and assess as many areas of potential policy impact as possible, to mitigate the risks that a given policy will have unexpected or unintended consequences.

Gamification of Poverty

The gamification module involves a reality-based poverty game system. It is envisioned to produce real-life game scenarios by using existing information contained in the poverty database and knowledge base. The script engine will generate stories and adapt game scenarios according to the participant's characteristics, sentiments, personal interests and community interests.

Figure 3
Game Participants



For poverty, the framework of Mechanics, Dynamics and Aesthetics (MDA) can be explored in the game design. As emphasized by Zicherman and Cunningham (2011), mechanics compose the functioning sections of the game that allow the game designer ultimate control over the levers of the game, enabling the ability to guide player actions while dynamics are the player's interactions with those mechanics. They determine what each player is doing in response to the mechanics of the system, both individually vis-à-vis other players. Aesthetics of the system are how the game makes the player feel during interaction. Game aesthetics can be regarded as the multifactorial outcome of the mechanics and dynamics as they relate with and generate emotions. The poverty gamification system can adopt game theory principles and multi-agent systems' concepts of collaboration, cooperation, and competition. We will treat every participant as an agent who engages collaboratively, cooperatively or competitively with other game participants, as shown in Figure 3, as they negotiate and achieve poverty reduction.

CONCLUSIONS

The use of gamification has always been associated with entertaining users by way of engaging them in games. However, gamification has been utilized in other countries as a behavior change mechanism particularly for people who are predisposed to playing games on the internet on issues that are reality-based. The change in behavior gets manifested in their opinions about an issue or problem and subsequently takes positive actions towards helping solve the problem. Researchers can explore to identify, via data mining techniques, relationships and trends from the CBMS database of household demographic and other relevant data for problem areas encountered by the community. Game design would be developed to help engage the members of the community to try to address different problem areas encountered.

For social relevance, many of the problems besetting poverty stricken households can be solved through change in behavior once they realize the opportunities that can come their way by

doing so. The change is much more effective if part of the motivation comes from within the individual itself and the immediate environment makes it possible for them to see the positive effect this change will have in their lives.

For the anticipated societal impact of implementation, the creation of more change agents in the community will greatly assist local government units in bringing into the community projects that could further improve the general well-being of everyone residing in it, particularly in the area of peace and order, sanitation and health and livelihood and employment.

Studies bringing technology very close to the grassroots level, particularly the poor and the marginalized sectors, are rare. Likewise, it seems to be the case that much of the developments seen with technology only seemingly address the needs of middle and upper income brackets of society. But the trickle-down effect can be hastened with studies of this nature so that hopefully other communities can learn from it. If technology can be introduced with ease that come with some form of entertainment to sustain interest in the challenge for all stakeholders, but keeping one eye to its ultimate poverty-alleviation purpose, data-collection is then done with ease, metrics-processing becomes efficient, results-generation more accurate and interventions more timely and effective.

The research design is highly experimental and would involve the following: (1) With the use of data mining techniques/procedures, identify trends and relationships between a community households' demographic make-up and the problems that beset the community; (2) Utilize such relationships to determine specific problem areas that can be addressed through community based action programs or policy reforms. If community based programs or policy reforms already exists, then there will be a need to enhance the effectiveness of delivery of service and dissemination of information on the benefits of availment of public service through behavior-changing games; (3) Survey on the extent of usage of internet cafes and cellphones by a sample of households within the community, to determine if this mode of information dissemination as well as data collection is feasible within the community. The lack of usage of this form of technology may warrant the need for alternative forms of information dissemination that can be made available in barangay or health centers assisted by personnel adept with the use of the technology. The surveys will also have to cover the area of literacy due to the need for this to make the technology usable to the community, particularly with reference to gamification. Participants will then have to be divided into two groups: those who can do the game by themselves and those who will have to be assisted; (4) A game is developed to address the issue affecting the community that would require behavior change. As in any gamification effort, the participant will have to be rewarded with some form of point system redeemable and exchanged into some basic need (like commodity) of the household. This would also entail the development of a reward system that would sustain the participants' interest in the game until the objective of the game is fulfilled; (5) Post-game evaluation on behavioral change is conducted to determine effectiveness of the game in addressing the community's problems and the modifications needed to improve the game in fulfilling its purpose if it can be used in other communities; (6) Consider the development of other games as may be deemed necessary to address other pressing issues within the sample community to sustain the success (if ever) and advantages put forth by previous gamification efforts; (7) Introduce a data collection mechanism for data collectors (researchers), possibly also in the form of a game, to be able to extend the reach of data collection to as many households. This would involve the

collection of quality data which is more accurate and which is updated constantly on a regular basis so that the community can easily move forth with programs attuned to immediate needs; and (8) A cost-benefit analysis is also to be conducted to compare the effectiveness of the intervention proposed and determine how the proposed intervention can complement existing services delivered in the community.

The prospects on behavior change that gamification brings can easily extend to as many areas that tend to plague informal dwellers and poverty stricken sectors that can assist in addressing health issues, fertility, sanitation, peace and order, employment, livelihood and subsequently, poverty reduction.

ENDNOTES

[1] This study was presented at the International Conference on Business, Economics, and Information Technology (ICBEIT) 2013 – *Doing Business in the Global Economy: Economic, Political, Social, Cultural and Technological Environments* held last 18-19 March 2013 at the Pacific International Hotel, Cairns, Queensland, Australia. **Disclaimer:** The findings, interpretations, and conclusions expressed in this work do not reflect the views of the author's institutional affiliations and its Board of Executive Directors. The contents of this work shall not be construed as a commitment by the author's institutional affiliation excet where provided for in a formal written agreement. The other usual disclaimer applies.

[2] *Remedios D. Bulos* is an Associate Professor at the Software Technology Department of the College of Computer Studies of De La Salle University. She obtained her Doctor of Philosophy in Computer Science and Artificial Intelligence from the University of Sussex – England; her Master of Economics with Specialization in Computer Science and Artificial Intelligence from Wakayama University - Japan; and her Bachelor of Science in Industrial Engineering from the University of Santo Tomas. Her research interests include Databases, Data Mining, Big Data Analytics, Business Intelligence, Artificial Intelligence, E-Commerce and Technology Entrepreneurship. **Email**: remedios.bulos@dlsu.edu.ph / remediosdedios@yahoo.com

[3] *Neriza M. Delfino* is an Assistant Professor at the Financial Management Department of the Ramon V. Del Rosario – College of Business of De La Salle University. She obtained her Doctor of Business Administration from De La Salle University; her Master of Business Administration from Katholieke Universiteit Te Leuven; and her Bachelor of Science in Applied Economics from De La Salle University. Her research interests include Economics, International Finance, and Microeconomics. **Email**: neriza.delfino@dlsu.edu.ph / iris_delfino@yahoo.com

[4] *John Paolo R. Rivera* is an Assistant Operations Manager at Veritas Financial Partners (VFP). Prior to joining VFP, he was an Associate Professor at the School of Economics of De La Salle University. He was also a Research Associate at the Angelo King Institute for Economic and Business Studies. He was the Managing Editor of the *DLSU Business & Economics Review*. He obtained his Doctor of Philosophy in Economics, Master of Science in Economics, and Bachelor of Science in Applied Economics (Magna cum laude) from De La Salle University. His fields include Macroeconomics and Macroeconomic Modeling, International Trade, Econometrics, Mathematical Economics, and

REFERENCES

Adeyemo, A. B., & Kuye, G. (2006). Mining students' academic performance using decision tree algorithms. *Journal of Information Technology Impact*, 6(3), 161-170.

Alexander, D. (n.d.). *Data mining*. Retrieved from: http://www.laits.utexas.edu/~anorman/BUS.FOR/course.mat/Alex/

Coren, M.J. (n.d.). *Live life below the poverty line by playing spent.* Retrieved from http://www.fastcompany.com/1784810/live-life-below-poverty-line-playing-spent

Han, J., Kamber, M., & Pei, J. (2011). *Data mining concepts and techniques* (3rd Edition). Masachusettes: Elsevier.

Herger, M. (2012). *Gamification Facts & Figures*. Retrieved from http://www.Enterprise-Gamification.com.

Koh, H.C., & Tan, G. (2005). Data mining applications in healthcare. *Journal of Healthcare Information Management , 19* (2), 64-72.

Lopez, J. (2011). *Weekly recap: Gamification around the web*. Retrieved from http://www.gamification.co/2011/10/07/weekly-recap-gamification-around-the-web-8/

McGonigal, J. (2010). *Jane McGonigal: Gaming can make a better world*. Retrieved from TED: http://www.ted.com/talks/jane_mcgonigal_gaming_can_make_a_better_world.html

Mims, C. (2011). *The gamification of Poverty*. Retrieved from http://www.technologyreview.com/view/425686/the-gamification-of-poverty

National Anti-Poverty Commission (NAPC). (2012). *Targeting the poor: A search for alternative methodologies.* Quezon City, Philippines: National Anti-Poverty Commission.

National Statistical Coordination Board (NSCB). Website: http://www.nscb.gov.ph.

Nicholson, J. (2012). *Spent*. Retrieved from http://playspent.org/.

Rivera, J.P.R., & See, K.G.T. (2011). The role of education and government sponsored programs in limiting family size in Pasay, Eastern Samar, and Agusan Del Sur. *DLSU Business & Economics Review*, 21(2), 17-36.

Rivera, J.PR., Pizarro, M.G., Aliping, N.B., & Reyes, P.O. (2012). **The implications of government's poverty reduction programs on the states of poverty and hunger in the Philippines. *Journal of International Business Research*, 11(2), 73-96.**

Scott, R. (2012, June 21). Gamifying employment engagement. *Business Fights Poverty* . Retrieved from http://www.businessfightspoverty.org/profile/blogs/gamefying-employee-engagement

Syed, A. (2011). *A research study on privacy in data mining*. Infia: Shri Jagdishprasad Jhabarmal Tibrewala University. Retrieved from: http://shodh.inflibnet.ac.in/handle/123456789/34

United Nations Development Programme (UNDP). (2012). *Poverty assessment and monitoring.* Retrieved from http://www.undp.org/content/undp/en/home/ourwork/povertyreduction/focus_areas/focus_poverty_assessment/

Zicherman, G., & Cunningham, C. (2011). *Gamification by design.* Sebastopol, Canada: O'Reilly.

THE IMPACT OF THE CARIBBEAN BASIN INITIATIVE PROGRAM ON THE ECONOMIC GROWTH & DEVELOPMENT IN THE ENGLISH SPEAKING CARIBBEAN REGION

Michael Campbell, Florida A&M University

ABSTRACT

The purpose of this paper is to examine the impact of the Caribbean Basin Initiative (CBI) Program on the English-speaking Caribbean countries and assess the impact on foreign direct investment (FDI), economic development and growth in the region. Export and import data between the United States and the CBI region were examined for the period 1994-2009 to determine the direction of trade. The Gross Domestic Product (GDP) of the CBI countries for the same periods was examined to determine if any economic growth and development had occurred. Over the period of this study, the data indicate that the Caribbean Islands, with the exception of Trinidad and Tobago, experienced negative balance of trade and deficit balance of payments. The results indicated that the CBI impact on the English- speaking Caribbean countries did not meet expectations as relate to economic development and growth.

Keywords: Caribbean Basin Initiative, English-speaking Caribbean countries, Economic Growth &Development

INTRODUCTION

The U.S. Congress enacted the Caribbean Basin Initiative *(CBI)* in 1984 to assist countries in Central America and the Caribbean Islands. The act was a linchpin in the U.S. effort to stabilize the Caribbean Basin during the 1980's. The principal economic objectives were to stimulate foreign and domestic investment, to diversify local economies, and to augment export earnings by eliminating U.S. customs duties on most items manufactured or assembled in the region. The CBI, first proposed in 1982, is a broad United States foreign policy designed program to promote economic development and political stability. The CBI is not limited to the Commonwealth Caribbean nations but extends to the entire Caribbean Basin, also including selected countries of Central America, northern South America, and the non-English-speaking Caribbean. The CBI consists of trade, economic assistance, and investment incentive measures to generate economic growth in the region through increased private sector foreign direct investment (FDI) and economic development (Lunger, 1987; Newfarmer, 1985).

The most significant aspect of the program is the Caribbean Basin Economic Recovery Act (CBERA) of 1983. The CBERA provide Caribbean Basin countries with duty-free access to the United States market for most categories of exported products until September 30, 1995. It also includes special tax provisions for the tourist sector, as well as measures to support the economic development of Puerto Rico and the United States Virgin Islands. In addition to the CBERA, other CBI measures include increased United States economic assistance, a wide range of government and private sector investment promotion programs, support from multilateral developing institutions and their donor nations, and Caribbean Basin country self-help efforts. The CBI

resulted from a series of 1981 meetings involving United States, Canadian, and Caribbean Basin officials. In a July 1981 meeting in Nassau, the United States special trade representative and the United States Secretary of State met with the foreign ministers of Canada, Mexico, and Venezuela. Each agreed to support a multilateral action program for the region, within which each country and dependent territory would develop its own programs. Multilateral and bilateral meetings were held between the members of the so-called Nassau group and representatives of the Caribbean Basin countries (Lunger, 1987; Newfarmer, 1985).

The CBI package announced by President Ronald Reagan in a February 1982 address before the Organization of American States (OAS) consisted of foreign assistance, a free trade arrangement, and tax incentives for United States investors. The foreign aid portion of the CBI, which proposed an additional U.S. $350 million for the Caribbean region for fiscal year 1982, was passed by the 97th Congress and became law in September 1982 (Two-thirds of this total was slated for Central America, with the remainder earmarked for the Caribbean.) (Zorn & Mayerson, 1983). The trade portion, contained in the CBERA, was passed by the 98th Congress in July 1983 and signed into law in August 1983 (Clasen, 1983). The CBERA also contained a tax benefit allowing United States citizens and companies to make deductions for expenses from conventions and business meetings held in CBI countries. The investment tax incentive portion of the package was left out of the legislation's final version. Also, a number of products were excluded from the eligibility list of duty-free exports (Newfarmer, 1985; Shingetomi, K., Rule, K., & Osler, D., 2009).

Twenty countries (20) and dependent territories were designated to receive benefits on January 1, 1984: Antigua and Barbuda, Barbados, Belize, British Virgin Islands, Costa Rica, Dominica, Dominican Republic, El Salvador, Grenada, Guatemala, Haiti, Honduras, Jamaica, Montserrat, Netherlands Antilles, Panama, St. Kitts and Nevis, Saint Lucia, Saint Vincent and the Grenadines, and Trinidad and Tobago. The Bahamas was designated on March 14, 1985. On April 11, 1986, Aruba was designated retroactively to January 1, 1986, upon becoming independent of the Netherlands Antilles. Guyana was designated effective November 24, 1988, and Nicaragua was designated effective November 13, 1990. This brought the total number of beneficiary countries to twenty-four. Anguilla, Cayman Islands, Suriname, and Turks and Caicos Islands have also been identified by Congress as potentially eligible for benefits, but have not yet requested beneficiary status (Shingetomi, K., Rule, K., & Osler, D., 2009). The English speaking Caribbean countries which are targeted in this study are as follows (by sub-region): (1) Leeward Islands: Antigua/Barbuda, St. Kitts/Nevis and Bahamas; (2) Windward Islands: Dominica, Grenada, St Lucia and St. Vincent; and (3) Greater Antilles: Barbados, Guyana, Jamaica and Trinidad & Tobago.

The purpose of this study is to examine the impact of the Caribbean Basin Initiative (CBI) Program on the English-speaking Caribbean countries in terms of economic development, and growth in the region.

REVIEW OF THE LITERATURE

The incoming Reagan administration in 1981 identified the Caribbean Basin as one of the most important regions in the world to the United States, and proposed a long-term economic program for the Caribbean designed to realize economic development and secure United States presence in the area (Sutton 1995). While the United States presence in the area was achieved, it did very little to advance the economic development of the region in any meaningful way. One reason for this is that the program was designed to reflect the political and security interests of the United States rather than those of Caribbean Basin countries. The CBI program did not provide

enough in the way of economic assistance. The trade concessions were minimal at best and the development incentives were compromised by an ideology that the private sector could be more successful than the public sector to achieve growth.

In addition, the program was not a matter of negotiation but rather that of imposition. In the best traditions of U.S. policies in the region, it was decided unilaterally in Washington and supported by reluctant allies, which included not only the countries of the region collectively, but also Britain and Canada. The effects of this on the policies of the metropolitan powers and the regional integration movement were examined by Leys, Hyett, and Moore, with a comparison of the Canadian aid program with the CBI program being particularly useful (Bakan, Cox & Leys, 1993).

A primary economic objective of the CBI is to assist Caribbean Basin countries earn their own way towards growth and economic development by providing aid to encourage private sector activities, by allowing one-way duty free exports to the U.S. and through the stimulation of private investments both at a national and foreign levels (Griffith, 1990). The one-way duty free trading with the United States only conferred marginal benefits on very small beneficiary CARICOM countries was definitely at variance with conventional economic theory which stated that when a very small country and a very large country integrate their economies by eliminating tariffs, the small country will appropriate maximum gains from the trade (Kenen, 1985). It is important to note that the free trade between the United States and the Caribbean countries, does not meet all the assumptions required for the application of the model. Conversely, from the available trade data, it is evident that the United States has gained more from the CBI agreement than the CARICOM countries (Zorn & Mayerson, 1983).

Another objective of the CBI program was to assist Caribbean Basin countries in stabilizing their economies through increased foreign direct investment which should diversify local economies and to augment export earnings through the elimination of tariffs (Woodward & Rolfe, 1993). Although the program has been in operation for almost three decades, the consensus by the CARICOM countries heads of governments is that the program had fallen extremely short of prior expectations. The political leadership strongly contended that the CBI program has been modest at best and have provided several reasons to support their arguments. The political leader of Dominica indicated in 1987, that the removal of US Congress of investment incentives from the original legislation discouraged US investors from locating in the CARICOM region thus depleting the effectiveness of the program foreign direct investment.

A member of the Jamaican political leadership argued that inconsistent and arbitrary rulings by the US Customs have negatively affected exports from CARICOM countries and resulted in the loss, due to non delivery of millions of dollars in orders from the region. Also, the restrictions imposed on the principal exports from the region were alluded to by CBI countries in order to explain and criticize the modest performance by the CBI program (Griffith, 1990). The major products excluded from duty-free entry were textiles and most leather goods. After the establishment of the CBI, direct investment in nontraditional products grew considerably. To evaluate the CBI, the U.S. Department of Commerce collected information on employment, location of ownership, value of investment, a company's markets, and product type from 642 companies.

The database encompasses local and nonlocal investments in agriculture, tourism, and manufacturing. It included new plant openings as well as plant expansions. Of the total 642 reported observations, only nonlocal (foreign) investments were analyzed in that study. Moreover, tourism-related and agricultural investments were excluded, leaving just manufacturing plants. Finally, only new plant investments were considered; expansions of existing facilities were omitted

leaving 187 observations (Woodard & Rolf, 1993). In response to the claims by the CARICOM political leadership, United States argued the CBI program was not designed to promote the chief exports of the region and that the Caribbean governments should adopt the necessary policy changes to improve the local, business climate (Good, 1988).

THEORETICAL FOUNDATION

The impact of trade on the economies of nations has often been aligned to the market model, which suggests that the best economic results are experienced by nations that adopt a policy of free flow of export and imports without trade restrictions. While linking trade to economic development, Grossman and Helpman (1990) concluded that nations practicing a free trade policy grew at a faster rate than those that did not practice such a policy. The researchers also concluded that developing nations stood to gain more from unrestricted trade, as in some instances developing nations do not possess the necessary resources to properly and effectively conduct the developmental research that is required for new product development. Grossman and Helpman (1990) confirmed the earlier research of Heller and Porter (1978) and Balassa (1978), which established that increased exports will successfully accelerate the economic growth and financial stability of the nations since exports are the main component of national outputs.

In order to quantify exports as a viable component of economic development, Tyler (1981) established that 17.5 percent increase in exports resulted in an incremental increase of one percent in GDP and that nations that do not conform to free trade policies will result in increased exports and economic growth being hampered. Thus, finding can be represented by the following equation:

$$17.5\% \, \Delta \text{ Exports} = 1\% \text{ GDP} \qquad (1)$$

Feder (1982), in an attempt to quantification, even separated the export from the non-export components of output and used the simple equation:

$$Y = N + X \qquad (2)$$

Where the GDP (Y), was equal to N + X, and N represented the non export sector and X the export sector. Therefore, it is fair to assume that regional integration and its export promotion policies are critical factors in advancing economic growth and development. Equation 2 may be rewritten as follows

$$\text{GDP} = \text{Non Exports} + \text{Exports} \qquad (3)$$

The balance of trade (BOT) is the difference between the monetary value of exports and imports in an economy over a certain period. It is the relationship between a nation's imports and exports. A positive or favorable balance of trade is known as a trade surplus if it consists of exporting more than is imported; a negative or unfavorable balance is referred to as a trade deficit or, informally, a trade gap. The balance of trade is sometimes divided into a goods and a services balance. The balance of trade is the difference between a nation's exports of goods and services and its imports of goods and services, if all financial transfers, investments and other components are ignored. A nation is said to have a trade deficit if it is importing more than it exports. BOT may be expressed by the equation:

$$BOT = EX - IM \quad (4)$$

Where BOT is the Balance of Trade; EX represents country's total exports; and IM represents country's total imports.

HYPOTHESIS:

The following hypotheses were developed and tested in this study:

Hypothesis 1: Duty free exports to the United States had an effect on economic growth, of the CBI countries' as measured by the GDP

Null 1: Duty free exports to the United States had no effect on economic growth, of the CBI countries' as measured by the GDP

*Hypothesis 2: **BOT** with the United States had an effect on economic development, of the CBI countries' as measured by the GDP*

*Null 2: **BOT** with the United States had no effect on economic development, of the CBI countries' as measured by the GDP*

RESEARCH METHOD

Sample and Data Collection

The sample included in this study consisted of **ten** Caribbean Basin countries: The English speaking Caribbean countries which are targeted in this study are as follows (by sub-region): (1) Leeward Islands: Antigua/Barbuda, St. Kitts/Nevis and Bahamas; (2) Windward Islands: Dominica, Grenada, St Lucia and St. Vincent; and (3) Greater Antilles: Barbados, Guyana, Jamaica and Trinidad & Tobago.

Export and import data, between the United States and the CBI region, as related to the CBI program, were first isolated from trade with other countries and then examined for the period 1994-2009. Data on trade balance, including imports and exports, were obtained from the International Monetary Funds (IFM) and the Direction of Trade Year Book for years 1994 to 2009.

Measurement of Variables

Gross domestic product (GDP) and trade balance, including imports and exports, for the CBI countries were measured as follows:

1. Gross domestic product (GDP): Gross domestic product data were calculated from the change in exports based on the formula introduced by Tyler (1981): 17.5% Δ Exports = 1% GDP.

2. Trade Balance data: Export and import data were required for the calculation of trade balance (BOT) of the CBI countries. Trade balance was measured based on the equation:

BOT = EX – IM.

Statistical Tools and Data Analysis

The Scientific Package of Social Sciences (SPSS) software was utilized to conduct the regression analysis. In order to ascertain the reliability of the constructs, Cronbach Alpha was calculated. Both the Greater Antillies and Leeward Islands Cronbach Alpha was 0.769 or 75% reliability. The Windward Islands Cronbach Alpha was -01.828. Factor Analysis was introduced to the data and the Descriptive Statistics were calculated, namely, Mean (m), Standard Deviation (s) and Dispersion (s^2) or (Variance).

The trade data was then analyzed to determine whether there were any increases in CBI exports to the United States. If increased trade occurred as a direct result of the elimination of tariffs and trade restrictions between United States and the CBI countries then the CBI program would have achieved its intent. The CBI countries export data to the United States was compared with the import data to determine the balance of trade amounts. It is important that the dollar amounts of CBI exports be greater than the CBI imports from the United States which will result in positive balance of trade. The gross domestic product (GDP) of the CBI countries that was generated for the same periods was examined to determine if any economic growth and development had occurred.

Increases in the CBI countries GDP as direct result of increases in exports to the United States will indicate growth and a stabilization of the CBI economies. The data was then subject to analysis utilizing SPSS software to calculate Mean, Standard Deviation (STD), Dispersion, and Analysis of Variance (ANOVA) for regression for each of the Caribbean regions participating in the CBI program. The export data of the CBI countries will be examined to determine and tabulate any increases in exports as these increase will only result from sustained both foreign and local direct investments in creating these exports. The continued increase in exports will result in an improving economy and increased economic development.

GROSS DOMESTIC PRODUCT, EXPORT AND IMPORT DATA CBI COUNTRIES

Export Data

The export data between the English speaking CBI countries and the United States of America for the period 1994 to 2009 are tabulated in *Table 1*. The data shows that overall the CBI region exports to the U.S. increased each year, from 1994 to 2008 by $8.759 billion, except for the period 2008 to 2009 where there was a massive decrease in exports to the U.S. by $3.920 billion.

The down turn of the U.S. economy was attributed as the reason for this decrease in exports to the U.S. The only segment of the CBI Region which did not experience decreased levels of exports to the U.S., was the Leeward Island region, which did experience increased exports in 2008 to 2009 of $219 million dollars.

This continued increase of exports by the Leeward Islands group was mainly due to the Bahamas which has continued growth in exports to the U.S. The largest exporter from the CBI Region was Trinidad & Tobago who had increased exports to the U.S. from 1994 to 2008 of $8.272 billion dollars mainly due to petroleum products and ethanol exports to the U.S. From 1994 to 2009, Jamaica experienced decline in exports to the U.S. which was attributed to the advent of the

North American Free Trade Agreement (NAFTA) between the U.S., Canada, and Mexico. With the implementation of NAFTA in 1994, many manufacturing companies moved their operations from Jamaica to Mexico in order to receive the 100% duty free status for their exports to the U.S. (Michaely, 1977; Newfarmer, 1985).

Table 1:

Total CBI Countries individual Exports to the United States

Caribbean Basin Initiative Regions

	Leeward Islands				Windward Islands					Greater Antilles					Region
Year	Total	Ant	St K	Bah	Total	Dom	Gre	St L	St V	Total	Bar	Jam	Guy	T&T	CBI
1994	247	5	22	220	49	7	8	28	6	2,144	36	790	119	1199	2,440
1995	198	3	24	171	57	7	6	36	8	2,130	52	895	129	1054	2,385
1996	212	9	25	178	43	9	4	23	7	2,168	43	890	129	1106	2,423
1997	217	5	32	180	57	10	7	35	5	2,181	43	780	131	1227	2,455
1998	191	2	35	154	48	7	13	23	5	2,061	37	798	155	1071	2,300
1999	241	2	39	200	79	23	20	28	8	2,341	60	728	146	1407	2,661
2000	319	2	39	278	68	8	27	24	9	3,223	41	669	160	2353	3,610
2001	376	4	44	328	85	6	25	31	23	3,287	41	495	161	2590	3,748
2002	533	4	51	478	53	9	7	20	17	3,255	36	421	134	2664	3,841
2003	562	14	48	500	32	6	8	14	4	5,414	45	524	136	4709	6,008
2004	716	5	44	667	28	4	5	15	4	6,783	38	341	138	6266	7,527
2005	787	5	56	726	60	4	6	34	16	8,919	33	411	133	8342	9,766
2006	538	6	57	475	42	3	5	32	2	9,560	35	562	141	8822	10,140
2007	593	9	61	523	48	2	9	36	1	10,318	40	789	147	9342	10,959
2008	693	5	61	627	39	3	8	27	1	10,467	42	784	170	9471	11,199
2009	912	10	57	845	29	3	6	19	1	6,338	34	501	179	5624	7,279

Source:Export data obtained from the International Monetary Fund (IMF) Direction of Trade Yearbook

Import Data

The import data between the English-speaking CBI countries and the U.S. is presented in *Table 2*. The CBI region did increase its imports from the U.S. for the period 1994 to 2008 by $6.464 billion dollars. For the period 2008 to 2009, there was a decrease in CBI imports by $2.089 billion dollars. The Greater Antilles, more than any of the other sub-regions, experienced the largest total imports from U.S. in the amount of $3.774 billion dollars from 1994 to 2008.

For the period 2008 to 2009 the Greater Antilles actually had a decrease in imports from the U.S. by $1.549 billion dollars. The Leeward Islands imports from the U.S. also increased substantially during the period, 1994 to 2009, by $2.345 billion dollars. This sub-region also experienced a decrease in imports for the period 2008 to 2009 by $0.375 billion dollars.

Table 2

Total CBI Countries individual Imports from the United States

Caribbean Basin Initiative Regions

	Leeward Islands				Windward Islands					Greater Antilles					Region
Year	Total	Ant	St K	Bah	Total	Dom	Gre	St La	St V	Total	Bar	Jam	Guy	T&T	CBI
1994	803	59	58	686	169	26	24	81	38	1878	161	1066	110	541	2850
1995	816	97	58	661	175	25	27	81	42	2452	201	1421	141	689	3443
1996	859	82	52	725	199	34	36	84	45	2655	362	1491	137	665	3713
1997	951	85	56	810	222	38	41	89	54	2947	281	1417	143	1106	4120
1998	973	96	62	815	490	52	56	93	289	2714	281	1304	146	983	4177
1999	1009	95	69	845	306	40	65	99	102	2562	300	1305	147	810	3877
2000	1273	137	83	1053	265	37	80	104	44	2932	309	1360	159	1104	4470
2001	1184	96	66	1022	220	31	60	89	40	2924	286	1407	141	1090	4328
2002	1127	82	70	975	240	45	57	98	40	2835	269	1420	128	1018	4202
2003	1292	127	81	1084	270	34	68	121	47	2952	302	1469	117	1064	4514
2004	1389	125	82	1182	254	36	70	103	45	3121	348	1460	136	1207	4764
2005	2085	190	126	1769	323	61	82	135	45	3694	393	1687	175	1439	6102
2006	2652	194	170	2288	354	68	76	152	58	4272	443	2035	179	1615	7278
2007	2916	240	203	2473	402	84	83	166	69	4742	457	2318	188	1779	8060
2008	3148	183	205	2760	514	105	85	241	83	5652	498	2644	259	2251	9314
2009	2773	157	161	2455	349	77	59	136	77	4103	405	1448	261	1989	7225

Source: Import data obtained from the International Monetary Fund (IMF) Direction of Trade Yearbook

Balance of Trade

The Balance of Trade of the English-speaking Caribbean countries with the U.S. were calculated and tabulated in Table 3. With the exception of Trinidad & Tobago which had positive balance of trade for each year of the period 1984 to 2009 and Guyana who had positive balance of trade figures for only seven years for the period, all the other countries experienced negative balance of trade. This means that they imported more goods from the U.S. than they actually were able to export. Therefore, with a negative balance of trade, it would have been highly impossible for the CBI countries, with the exception of Trinidad and Tobago, to have achieved economic development and growth (Michaely, 1977).

Gross Domestic Product (GDP) did not indicate a positive significant correlation with any of the sub-regions. The only significant correlation was the Greater Antilles with a significant negative correlation with imports from the U.S. This result was very surprising as the GDP was calculated from the CBI countries exports to the U.S based on the basis which Tyler (1981) established that 17.5 percent increase in exports resulted in an incremental increase of one percent in GDP. As relates to the Leeward Islands sub-region, there was significant correlation between exports and imports of 0.765. Also there was a significant positive correlation between exports and imports of 0.975. Surprisingly, the correlation between balance of trade and exports was significantly negative in the amount of -0.607. As balance of trade was calculated as the net of exports and imports for the sub-region, it was expected that there would be a positive correlation between balance and trade and the exports data. As relates to the Greater Antilles, there were significant positive correlations between balance of trade and imports and exports of 0.776 and 0.939 respectively.

Table 3															
CBI Balance of Trade Data with the United States of America															
Caribbean Basin Initiative Regions															
	Leeward Islands				Windward Islands					Greater Antilles					Region
Year	Total	Ant	St K	Bah	Total	Dom	Gre	St Lu	St V	Total	Bar	Jam	Guy	T&T	CBI
1994	(542)	(54)	(22)	(466)	(120)	(19)	(16)	(53)	(32)	266	(125)	(276)	9	658	(396)
1995	(590)	(94)	(6)	(490)	(118)	(18)	(21)	(45)	(34)	(332)	(149)	(526)	(12)	365	(1,040)
1996	(629)	(73)	(9)	(547)	(156)	(25)	(32)	(61)	(38)	(487)	(319)	(601)	(8)	441	(1,272)
1997	(723)	(80)	(13)	(630)	(165)	(28)	(34)	(54)	(49)	(766)	(238)	(637)	(12)	121	(1,654)
1998	(780)	(94)	(25)	(661)	(442)	(45)	(43)	(70)	(284)	(653)	(244)	(506)	9	88	(1,875)
1999	(747)	(93)	(9)	(645)	(227)	(17)	(45)	(71)	(94)	(221)	(240)	(577)	(1)	597	(1,195)
2000	(952)	(135)	(42)	(775)	(197)	(29)	(53)	(80)	(35)	3,291	(268)	(691)	1	4,249	2,142
2001	(811)	(92)	(25)	(694)	(135)	(25)	(35)	(58)	(17)	363	(245)	(912)	20	1,500	(583)
2002	(609)	(78)	(34)	(497)	(187)	(36)	(50)	(78)	(23)	420	(233)	(999)	6	1,646	(376)
2003	(733)	(113)	(36)	(584)	(238)	(28)	(60)	(107)	(43)	2,462	(257)	(945)	19	3,645	1,491
2004	(679)	(120)	(44)	(515)	(226)	(32)	(65)	(88)	(41)	3,632	(310)	(1,119)	2	5,059	2,727
2005	(1,321)	(185)	(93)	(1,043)	(263)	(57)	(76)	(101)	(29)	5,225	(360)	(1,276)	(42)	6,903	3,641
2006	(2,136)	(188)	(135)	(1,813)	(312)	(65)	(71)	(120)	(56)	5,288	(408)	(1,473)	(38)	7,207	2,840
2007	(2,344)	(231)	(163)	(1,950)	(354)	(82)	(74)	(130)	(68)	5,576	(417)	(1,529)	(41)	7,563	2,878
2008	(2,474)	(178)	(163)	(2,133)	(102)	(102)	(77)	(214)	(82)	4,815	(456)	(1,860)	(89)	7,220	2,239
2009	(1,884)	(147)	(127)	(1,610)	(320)	(74)	(53)	(117)	(76)	2,253	(371)	(947)	(82)	3,653	49
Source:Balance of Trade data calculated as the Net of Exports and Imports with the USA															

Gross Domestic Product (GDP)

The Gross Domestic Product for each CBI country was calculated on the basis which Tyler (1981) established that 17.5 percent increase in exports resulted in an incremental increase of one percent in GDP. This finding can be represented by the following equation:

$$17.5\% \, \Delta \text{ Exports} = 1\% \text{ GDP}$$

The calculated GDP amounts for each country is posted in Table 4 where it is very visible that each of the English-speaking Caribbean sub-region either experienced little growth, zero growth or negative growth during the period from 1994 to 2009. The Greater Antilles sub-region experienced the largest growth during the period mainly due to Trinidad & Tobago increased export of petroleum products and natural gas to the U.S. However the level of growth by the country Trinidad & Tobago was not consistent and in years 2008 and 2009 experienced large decreases in the gross domestic product.

Table 4

GDP Calculations Based on CBI Countries individual Exports to the United States.

Caribbean Basin Initiative Regions

	Leeward Islands				Windward Islands					Greater Antilles				
Year	Total	Ant	St K	Bahs	Total	Dom	Gre	St L	St V	Total	Bar	Jam	Guy	T&T
1994	(2.80)	(0.01)	(2.80)	0.10	0.50	0.00	(0.10)	0.50	0.10	(0.80)	0.90	6.00	0.60	(8.30)
1995	0.80	0.30	0.40	0.10	(0.80)	0.10	(0.10)	(0.70)	(0.10)	2.20	(0.50)	(0.30)	0.00	3.00
1996	0.30	(0.20)	0.10	0.40	0.90	0.10	0.20	0.70	(0.10)	0.70	0.00	(6.30)	0.10	6.90
1997	(1.50)	(0.20)	(1.50)	0.20	(0.60)	(0.20)	0.30	(0.70)	0.00	(6.80)	(0.30)	1.00	1.40	(8.90)
1998	2.90	0.00	2.60	0.20	1.80	0.90	0.40	0.30	0.20	16.00	1.30	(4.00)	(0.50)	19.20
1999	4.50	0.00	4.50	0.00	(0.60)	(0.90)	0.40	(0.20)	0.10	50.40	(1.10)	(3.40)	0.80	54.10
2000	3.30	0.10	2.90	0.30	1.00	(0.10)	(0.10)	0.40	0.80	3.70	0.00	(9.90)	0.10	13.50
2001	9.00	0.00	8.60	0.40	(1.70)	0.20	(1.00)	(0.60)	(0.30)	(1.80)	(0.30)	(4.20)	(1.50)	4.20
2002	1.70	0.60	1.30	(0.20)	(1.10)	(0.20)	0.10	(0.30)	(0.70)	123.40	0.50	5.90	0.10	116.90
2003	8.80	(0.50)	9.50	(0.20)	(0.20)	(0.10)	(0.20)	0.10	0.00	78.20	(0.40)	(10.50)	0.10	89.00
2004	4.10	0.00	3.40	0.70	1.90	0.00	0.10	1.10	0.70	122.00	(0.30)	4.00	(0.30)	118.60
2005	(14.10)	0.10	(14.30)	0.10	(1.10)	(0.10)	(0.10)	(0.10)	(0.80)	36.60	0.10	8.60	0.50	27.40
2006	3.10	0.20	2.70	0.20	0.20	(0.10)	0.20	0.20	(0.10)	43.30	0.30	13.00	0.30	29.70
2007	5.70	(0.20)	5.90	0.00	(0.50)	0.10	(0.10)	(0.50)	0.00	8.50	0.10	(0.30)	1.30	7.40
2008	12.60	0.30	12.50	(0.20)	(0.60)	0.00	(0.10)	(0.50)	0.00	(236.00)	(0.50)	(16.20)	0.50	(219.80)
2009	52.10	(0.60)	(48.30)	(3.30)	(1.70)	(0.20)	(0.30)	(1.10)	(0.10)	(362.10)	(1.90)	(28.60)	(10.20)	(321.40)

Source:GDP data obtained from the International Monetary Fund (IMF) Direction of Trade Yearbook

The other countries in the English-speaking Caribbean region experienced, on average from 1984 to 2009, negative growth. These below par GDP figures clearly indicate that the CBI Program did not have a positive impact on the English-speaking Caribbean countries.

FINDINGS OF THIS STUDY

Descriptive Statistics and Correlations

Table 5 shows the descriptive statistics and correlations. There was a significant correlation between exports to and imports from the U.S. of 0.897. The most astonishing of the correlation results was that gross domestic product, in all three sub-regions, was not positively significantly associated with exports from the CBI region, although the gross domestic product was actually calculated from exports to the U.S. For both the Leeward Island (0.765) and Greater Antilles (0.776) sub-regions, there was significant association between exports and imports. This result is in keeping with the reality that additional raw materials may have to be imported in order to sustain the increased levels of exports.

Table 5

Descriptive Statistics & Correlations

Windward Islands

	Mean	STD	Dispersion	1	2	3	4
Imports	297.00	103.71	10,756.67	1	-0.257	-0.583	0.001
Exports	51.06	16.47	271.40	-0.257	1	0.26	0.294
BOT	(222.63)	95.97	9,209.58	*-0.583	0.26	1	0.173
GDP	(0.84)	1.10	1.20	0.001	-0.294	-0.173	1

* Correlation is significant at the 0.05 level (2-tailed)

Leeward Islands

	Mean	STD	Dispersion	1	2	3	4
Imports	1,578.13	834.43	696,267.85	1	0.765	0.437	0.437
Exports	458.44	238.38	56,823.33	**0.765	1	0.464	0.464
BOT	(1,122.13)	682.41	465,683.45	**0.976	*-0.607	1	0.381
GDP	5.66	13.74	188.71	0.437	0.464	-0.381	1

** Correlation is significant at the 0.01 level (2-tailed)

* Correlation is significant at the 0.05 level (2-tailed)

Greater Antilles

	Mean	STD	Dispersion	1	2	3	4
Imports	3,277.19	971.73	944,253.36	1	0.897	0.776	0.519
Exports	5,036.81	3,231.58	10,443,076.98	**0.897	1	**0.939	0.254
BOT	1,945.75	2,393.22	5,727,492.40	**0.776	**0.939	1	0.128
GDP	(7.66)	123.36	15,216.89	*-0.519	-0.254	-0.128	1

** Correlation is significant at the 0.01 level (2-tailed)

* Correlation is significant at the 0.05 level (2-tailed)

HYPOTHESES TESTING

Finally, the results of the Analysis of Variance (ANOVA) as tabulated in *table 6,* show the F ratio for the subregions to be as follows: for the Leeward Islands sub-region, F-ratio = 1.948; for the Windward Islands subregion F-ratio = 0.555; and for the Greater Antilles subregion, F-ratio = 3.959. These results are not significant, indicating that BOT and exports had no effect on GDP.

As a result the Null Hypotheses 1 & 2 should be accepted thereby confirming that the Caribbean Basin Initiative did not have a positive impact on the English Speaking Caribbean Region as it relates to the regions, economic growth and economic development.

Table 6						
Analysis Of Variance (ANOVA)(b)						
Leeward Islands						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	652.629	2	326.314	1.948	0.182(a)
	Residual	2178.07	13	167.544		
	Total	2830.699	15			
a. Predictors: BOT, Exports, Imports b. Dependent Variable: GDP c. Countries: Antigua, St Kitts and Bahamas						
Windward Islands						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2.199	3	0.733	0.555	0.655(a)
	Residual	15.859	12	1.322		
	Total	18.058	15			
a. Predictors: BOT, Exports, Imports b. Dependent Variable: GDP c. Countries:Dominica, Grenada, St Lucia, St Vincent						
Greater Antilles						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	113470.801	3	37823.6	3.959	0.360(a)
	Residual	114782.518	12	9565.21		
	Total	228253.319	15			
a. Predictors: BOT, Exports, Imports b. Dependent Variable: GDP c. Countries: Barbados, Jamaica, Guyana, Trinidad						

CONCLUDING REMARKS

The U.S. Congress enacted the Caribbean Basin Initiative program in 1984 to assist Central America and the Caribbean region. The program was the main strategy of the U.S. effort to stabilize the Caribbean Basin. The principal economic objectives were to stimulate foreign investment, to diversify local economies, and to augment exports to the U.S. and its earnings by eliminating U.S. customs duties on most products manufactured or assembled in the region (Clasen, 1983, Woodward & Rolfe 1993). Although the ultimate impact of the CBI program on the region's economic development is subject to debate, the rise of foreign direct investment occurred in the early stages of the program implementation where foreign investments grew by almost $2 billion (Woodward & Rolfe, 1993). However, this growth of foreign investments was not continuous, but short lived. By 1986 to 1990, foreign investment inflows declined by 50% to less than $1 billion (Watson 1991, Woodward & Rolfe 1993). It was clearly indicated by both the Prime Ministers of Jamaica and Trinidad & Tobago, that unless there is sustainable foreign direct investment in the CBI Caribbean Region, the program only had a limited chance of success (Kenen, 1985; Watson, 1985; World Bank, 1985).

Research has indicated that when small countries, like those in the CBI region, and a very large country, like U.S., integrate by the elimination of tariffs, it is expected that the smaller countries will eventually accumulate gains from the trade (Kenen, 1985). Although the CBI program does not meet nor satisfy all the assumptions and requirements of the model as stated by Kenen (1985), it is extremely important to note U.S. seemed to be benefitting more from the

program than the CBI Caribbean countries (Review the Balance of Trade Table 4). In 1987, U.S. exports to the CBI Caribbean region increased by almost 18% over the 1983 data. However, for that same period, imports from the CBI region declined (Griffith, 1990).

The Prime Minister of Dominica, like her counterparts from Jamaica and Trinidad & Tobago, stated that the removal by U.S. Congress of the investment incentives from the original CBI program, discouraged U.S. investors from locating their operations in the CBI Region (Griffith, 1990; Tucker, 1987). In response to these allegations by the Caribbean Prime Ministers, the CBI Ombudsman argued that the intent of the CBI program was not to promote the major exports of the region and that it was the responsibility of the Caribbean governments to take that initiative and adopt the necessary policy changes to achieve that goal, along with improving the Caribbean countries infrastructure, as with little infrastructure, few export oriented investors will take an interest in the region (Tucker, 1987).

The results indicated that it is irrefutable that the CBI Region, with the exception of Trinidad and Tobago (an exporter of petroleum and natural gas products to the U.S.), experienced negative balance of trade with the U.S. The results further indicated that the CBI program impact on the English- speaking Caribbean countries did not meet expectations as relate to economic development and growth. The removal of foreign direct investment from the program by the U.S. Congress resulted in a negative impact on the region's economy development and growth.

REFERENCES

Bakan, A, Cox, D, & Leys, C., (1993). Imperical Power and Regional Trade: The Caribbean Basin Initiative. Waterloo: Wilfrid Laurier University Press.

Balassa, B. (1978). Exports and Economic Growth: Further Evidence, *Journal of Developmental Economics*, 5: 181-189.

Clasen, T. (1983). The Caribbean Basin Economic Recovery Act and Its Implications for Foreign Private Investment. *Journal of International Law and Politics.*16, 715 - 748.

Feder, G. (1982). On Exports and Economic Growth. *Journal of Development Economics*, 12:59-73.

Good, A. (1988). Progress in the Caribbean Basin: A Special Report by the CBI Ombudsman. *Caribbean Affairs,* 1, 1:16-28.

Griffith, W. (1990). Caribbean Crisis and Global Restructuring, *Latin American Perspectives*, 17,1:33-54.

Grossman, G. & Helpman, E. (1990). The New Growth Theory; Trade Innovation and Growth, *American Economic Association Papers and Proceedings, 86-89.*

Heller, P. & Porter R. (1978). Exports & Growth: An Empirical Reinvestigation, *Journal Developmental Economics, pp* 191-193.

International Monetary Fund (IMF), *Balance of Trade Statistics Yearbook*, 1995, 2000, & 2010.

Kenen, P. (1985). *The International Economy*. Englewood, Cliffs, NJ: Prentice-Hall.

Lunger, A., (1987). The Caribbean Basin Initiative and I.R.C. Section 936 Investment Program: A United States Answer to the Troubled Caribbean. *Journal of International Business Law,* 9, 4: 741 – 785.

Michaely, M. (1977). Exports and Growth: An Empirical Investigation, *Journal of Development Economics*, 4:49-53.

Newfarmer, R., (Feb., 1985). Economic Policy toward the Caribbean Basic: The Balance Sheet; *Journal of Interamerican Studies and World Affairs*; 27, 1: 63-89.

Shingetomi, K., Rule, K., & Osler, D. (December 31, 2009). Eight Report to Congress on the Operation of the Caribbean basin Economic Recovery Act. *Office of the United States Trade Representative, Office of the Americas. Executive office of the President.* Washington, D.C.

Tucker, S. (1987), Trade Unshackled: Assessing the Value of the Caribbean Basin Initiative. Paper prepared for *the International Commission on Central American Recovery and Development*. November.

Tyler, W. (1981). Growth and Export Expansion in Developing Countries, *Journal of Developmental Economics*, 9: 121-130.

Watson, H. (1985). The Caribbean Basin Initiative and Caribbean Development: A Critical Analysis. *Contemporary Marxism*, 10 (Summer): 1-37.

Watson, H., (1991). Coalition Security Development: Military Industrial Restructuring in the United States and Military Electronics Production in the Caribbean. *Caribbean Studies*. 24, 2/2: 223 – 247.

Woodard, P. & Rolfe, R., (1993); The Location of Export Oriented Foreign Direct Investment in the Caribbean Basin. *Journal of International Business Studies*, 34, 1: 121-144.

World Bank (1985). *St. Lucia Economic Performance and Prospects*. Washington, D.C.

Zorn, J., and Mayerson, H., (Winter 1983). The Caribbean Basin Initiative: A Windfall for Private Sector, *Lawyer of the Americas*, 14, 3: 523-556.

CROP YIELD PREDICTION USING TIME SERIES MODELS

Askar Choudhury, Illinois State University
James Jones, Illinois State University

ABSTRACT

Climate and other environmental changes in the developing world and the African continent has become a major threat to their agricultural economy. Traditional insurance for financial risk management is impractical in developing countries because of high transaction costs, adverse selection, information asymmetry, poor distribution and other challenges which hinder the availability of protection. Area-based index insurance is viewed as a promising financial risk management solution for smallholder farmers in developing countries, such as, Ghana. However, estimating the yield (i.e., yield prediction) is a critical part of pricing the premium for this insurance instrument. Because of the importance of predicting crop yield, the purpose of this study is to apply several forecasting methods for evaluating crop yield estimates in Ghana. Crop yield forecasting, which provides information for decision makers, is important in many ways to Ghana's economy. We compare yield forecasts using Simple Exponential Smoothing, Double Exponential Smoothing, Damped-Trend Linear Exponential Smoothing, and ARMA models applied separately to each district. The ARMA models proved to be more robust time-series models than the smoothing techniques for predicting crop yield in this study. This predictive power of ARMA models even with the presence of crop yield "cycle" does not depend on the length of cycle. Therefore, the results of this study indicate that the ARMA model is preferable over other time series models considered in this paper. The implication of the findings in this study is significant for insurance underwriters responsible for constructing area-based yield insurance that can benefit the Microinsurance market of smallholder farmers and for institutions that rely on those forecasts in providing capital.

INTRODUCTION

Farming is a major source of income for many people in developing countries. In Ghana farming represents 36 percent of the country's GDP and is the main source of income for 60 percent of the population (http://earthtrends.wri.org, 2003 p. 1). In addition, agricultural growth in Ghana has been more rapid than growth in the non-agricultural sectors in recent years, expanding by an average annual rate of 5.5 percent, compared to 5.2 percent for the economy as a whole (Bogetic et al., 2007). As with other parts of the developing world and the African continent, climate and other environmental changes in Ghana has become a major threat to their agricultural economy (Etwire et al., 2013). Direct losses to farming include destruction of their assets (such as, crop, livestock) which push poor farmers into poverty traps from which they have little means of recovery. Indirect impacts include sub-optimal management of this financial risk exposure, for example by selecting low-risk, low-return asset and activity portfolios that reduce the risk of

greater suffering, but limit growth potential and investment incentives, selling assets (at inopportune times), reducing nutrient intake, and withdrawing kids from school and hiring them out to work. The problem is exacerbated by the reaction of financial institutions, which may restrict lending to farmers to minimize exposure to agricultural risk. These indirect consequences hinder economic growth (Barnett et al., 2008).

Traditional insurance is impractical in developing countries because of high transaction costs, adverse selection, information asymmetry, poor distribution, and other challenges which hinder the availability of protection (Skees, 2008). Furthermore, post-event response in the form of emergency aid, debt forgiveness, and grants are at risk following recent economic crises, and such public capital does not usually help create independent private solutions and can be inequitable and untimely. In recent years, index based insurance instruments have been piloted as a way for smallholder farmers to hedge their losses. Unlike traditional indemnity insurance, the payout on index insurance products is not based on actual farm level yield and/or revenue losses. It is rather based on realizations of an index which assumes correlations with actual farm yield (or revenue) losses. Since the indexes are based on objective and transparent sources of data, it is unlikely that informational asymmetries exist that can be exploited by index insurance contract purchasers. Thus, the inherent insurance problems of adverse selection and moral hazard, additionally the high transaction costs of implementation can be largely avoided (Deng et al. 2006). Index insurance may also have the benefit of crowding-in capital, and allow farmers to get loans for needed inputs, as the risk for agricultural losses and thus financial risk becomes more manageable (Carter et al. 2007).

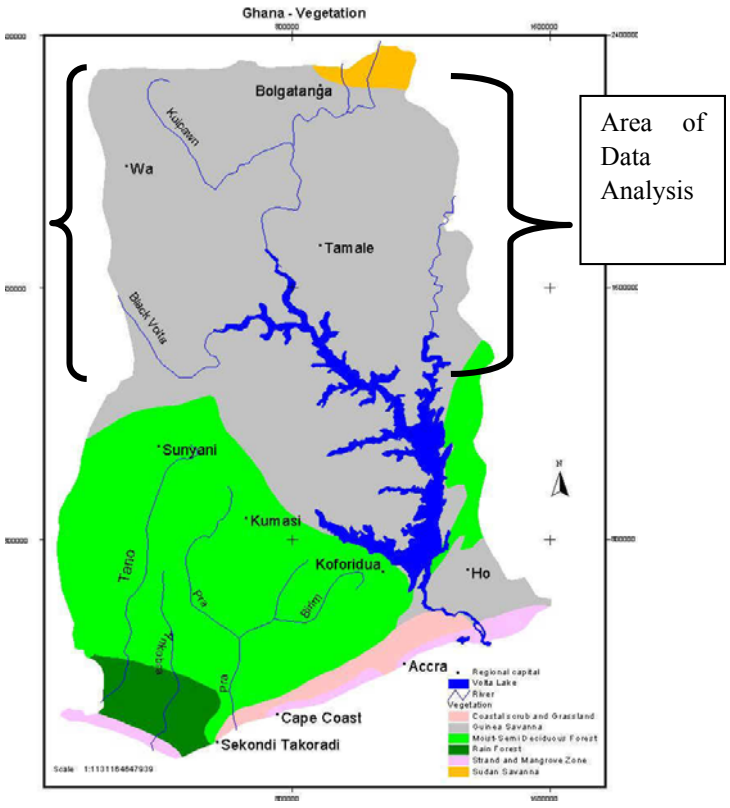
The two types of index products are parametric and sample-based. Examples of parametric indices in insurance include weather (with triggers based on variables such as rainfall, temperature, humidity, wind speed, etc.), flooding (water levels and durations triggers), wind speed (velocity and duration triggers) and seismic activity (Richter scale triggers). Sample based indices include area based yield insurance and sample based livestock index insurance. Area yield insurance is essentially a put option on the average yield for a production in a region/area. Payouts are triggered by shortfalls in that area average yield rather than farm level yield. For this reason, area yield insurance requires no farm-level risk underwriting or loss assessment. If the area is sufficiently large, area yield insurance is not susceptible to moral hazard problems, since the actions of an individual farmer will have no noticeable impact on the area average yield. Area yield insurance also has relatively low transaction costs since there is no need to establish and verify specific farm yields for each insured unit nor is there any need to conduct on-farm loss adjustment.

Crop yield (in Africa and many other countries) is defined as metric tons of production per hectare or area cropped. For an area, such as a district in Ghana, the calculation requires a sampling of production for a crop for the entire district and dividing by the area cropped in that district for that given crop. Ghana's Ministry of Food and Agriculture (MoFA) conducts sample crop-cutting at a district level for maize, rice and other food crops throughout Ghana and reports their results to the Statistics, Research and Information Directorate (SRID) in Ghana (Stutley. 2010). The reliability of crop cutting, in developing countries, is sometimes questioned because of variations in resources and expertise available. More reliable resources and more accurate sampling techniques, video recording crop cutting experiments with GPS-enabled cell phones and

remote sensing via satellite imagery, are new ways being piloted to help inform area yield estimations and make them more reliable.

Selection of Crop Region

Ghana produces a variety of crops in various climatic zones which range from dry savanna to wet forest. Agricultural crops including yams, grains, cocoa, oil palms, groundnuts and timber form the base of Ghana's economy. This research is focused mainly on the northern part of Ghana where there is substantial farming activity. The northern region of Ghana is considered the major bread basket of the country, and is also the most susceptible to the vagaries of the weather, especially the lack of rainfall. The northern part of Ghana is made up of three main regions; the Upper West Region, the Upper East Region and the Northern Region. The largest of these is the Northern Region which incidentally is the largest region in Ghana, covering a land area of about 70,383 square kilometers. However, it has the lowest population density of all the ten regions in the country (PPMED, Ghana, 1991) with 80% of its people dependent on farming. The major food crops grown here are yam, millet, rice, maize, sorghum, soybeans, groundnut and cassava.



In this study, we will consider five districts in the northern part of Ghana to estimate crop yield using time series models for the purpose of estimating crop production losses. Crops in this area are almost 100 percent rain fed (Stutley, 2008). Ghana is a country that is politically stable, has relatively good data and favorable regulation. A well designed financial risk management system in the agricultural sector could allow Ghana to act as a gateway to Africa for insurance underwriters who are not currently participating in Africa. As foreign donors have become increasingly diligent in assessing the need for loans and emergency relief, a credible index tied to true economic loss could be used by Ghana in justifying the need for emergency aid, loans and debt relief.

Accurate knowledge of crop yield behavior of the region is critical for devising such type of crop insurance product. Knowledge of the likelihood of yield and severity of yield shortfalls of the area are necessary components to create appropriate crop insurance. However, crop yield can be extremely dispersed from year to year and create complex scenario for predictability. Although understanding the stochastic nature of crop yield is important, characterizing yield behavior can be quite difficult. In general, historical yield distributions are used to set crop insurance premiums

based upon the assumption that the following year's realization is drawn from the same distribution.

Because of the importance of crop yield prediction, the purpose of this study is to apply several forecasting methods for evaluating crop yield forecasting models. Crop yield prediction, which provides information to decision makers, is important in many ways to the economy. Because of its importance, researchers have proposed many forecasting methods to improve accuracy of yield estimates. However, obtaining accuracy is not an easy task, as many factors have impacts on crop production and thus crop yield. Many methods have been used in yield forecasts and different models have generated different results. The most widely used is the Box-Jenkins ARMA (autoregressive moving average) models. ARMA models have been used to forecast maize production in Nigeria (Badmus and Ariyo, 2011), wheat production in Pakistan (Najeeb et al., 2005), rice production in Ghana (Suleman and Sarpong, 2012) and rainfall in Ethiopia (Gerretsadikan and Sharma, 2011). As accuracy and simplicity is a big concern in projection, researchers have begun to explore other methods in their forecasting. These include Simple Exponential Smoothing (Boken 2000; Pal et al. 2007), Double Exponential Smoothing (Boken 2000; Pal et al. 2007), and Damped-Trend Linear Exponential Smoothing. These predictive models can be ranked by R-square and other model performance criteria. This method of model evaluation is then applied to five widely used time series models implemented in this paper. We find ARMA (autoregressive moving average) method outperform the competing methods in predicting crop yields in all five districts considered in this study.

Variations from the predicted farm-level yields are largely a function of systemic risk such as the pervasive drought or excessive rain (Halcrow, 1949). An area yield policy has an associated basis risk when farmers' experience farm-level yield losses while the area yield shortfalls are not sufficient to trigger a payout under an area yield policy. This occurs when shock losses are idiosyncratic. Area yield insurance provides more effective risk management where yield risks are largely systemic. Lowering the chances of such an event (i.e., lowering the basis risk) is an important objective when designing an area yield insurance policy. The magnitude of the basis risk is affected primarily by two elements of the contract design: (a) the area to be used for the yield index and (b) the procedures for forecasting the yields for the area (Skees et al., 1997). Crop yield distribution primarily consists of average yield and standard deviation of yields. We expect average yield to stay same over time if the factors that influence the yield also move in tandem. Similarly, variations in yield would be similar also if the factors themselves affecting the yield stay same. However, extreme changes in those factors, such as, weather (e.g., drought, flood, hail, etc.) can influence the crop yield adversely and widen the yield variance. Therefore, the purpose of crop insurance is to provide protection against yield shortfalls due to these natural hazards. Thus, a prediction model to estimate the crop yield that accounts for higher percentage of yield variations is a preferable estimation model.

DATA AND RESEARCH METHODOLOGY

Data was collected from The Ministry of Food & Agriculture, which is the main government organization responsible for formulating and implementing agricultural policy in Ghana. The Statistics, Research and Information Directorate (SRID) and Policy Planning

Monitoring and Evaluation Division (PPMED) are two of the five directorates through which the ministry carries out its functions. According to information on the Ministry's website, the SRID has as some of its objectives "to initiate and formulate relevant policies/programs for creation of timely, accurate and relevant agricultural statistical database to support decision making" and "to conduct agricultural surveys and censuses covering major agricultural commodities". The PPMED, on the other hand, is responsible for undertaking, monitoring and evaluation of programs and projects under the Ministry. The statistical service department is an independent government department that is responsible for the collection, compilation, analysis, publication and dissemination of official statistics in Ghana for general and administrative purposes.

Crops which are likely to be suitable for Area-Yield Index Insurance include rain-fed maize and rice, and possibly millet, sorghum and groundnuts. This paper attempts to estimate the area "yield" of one crop, maize, for the purpose of creating an area-based index insurance instrument. Crop yield forecasting is primarily done with crop simulation models and empirical statistical regression equations relating yield with relevant predictor variables. These associative models require future data on the predictor variables. Crop forecasts are typically needed between the time of planting and the time of harvest. These associative models use past data to estimate the models and "future" data for prediction. Future data can be implicit or explicit. In general, forecasting methods can be subdivided into two categories: qualitative and quantitative (Makridakis et al., 1998; Armstrong, 2001) methods. Some of them are subjective, based on stakeholders' intentions or on the forecaster's or other experts' opinions or intentions, and others are objective/statistical, including univariate (extrapolation method), multivariate (associative method) and theory based methods. Other types may include expert systems or neural net, basically a variant of extrapolation with some subjective expert opinion.

Limitations of soil, weather and other relevant data cause a considerable uncertainty in the large area yield forecasting models (Hoogenboom, 2000; Russel and Gardingen, 1997). It is often unclear how these uncertainties transmit through the system given the non-linear behavior of crop yield models and the aggregation errors that may creep in when aggregating crop yields to larger regions (Hansen and Jones, 2000). Considerable amount of research to understand the effects of uncertainty in weather and other relevant factors on crop yield has been carried out by researchers. Crop yield modeling researchers primarily focused on local scale analyses in order to assess uncertainty in yield management (Bouman, 1994), condition of the soil (Pachepsky and Acock, 1998; Launay and Guérif, 2003), and weather components that affect crop yield (Fodor and Kovacs, 2005; Nonhebel, 1994; Soltani et al., 2004). In general, these studies demonstrate that the uncertainty in the modeling process is primarily a result of uncertainties in soil conditions and/or weather components. However, the local scale representation of these studies make the results less representative of regional scale crop yield forecast. Much of the research by climate researchers has been devoted to quantifying the climate variation effect on crop yield and studying the response of crop models to the climate change scenarios that are derived from general circulation models (GCMs). These research studies reveal that crop yield models are sensitive to the inconsistency of precipitation and temperature (Mearns et al., 2001; Semenov and Porter, 1995) and that the spatial scale of weather variables are also critical (Carbone et al., 2003; Mearns et al., 1999) to the crop yield prediction. In addition, when aggregating the yield at the regional scale, weather usually

becomes the primary uncertainty factor compared to the soil (Easterling et al., 1998; Mearns et al., 2001).

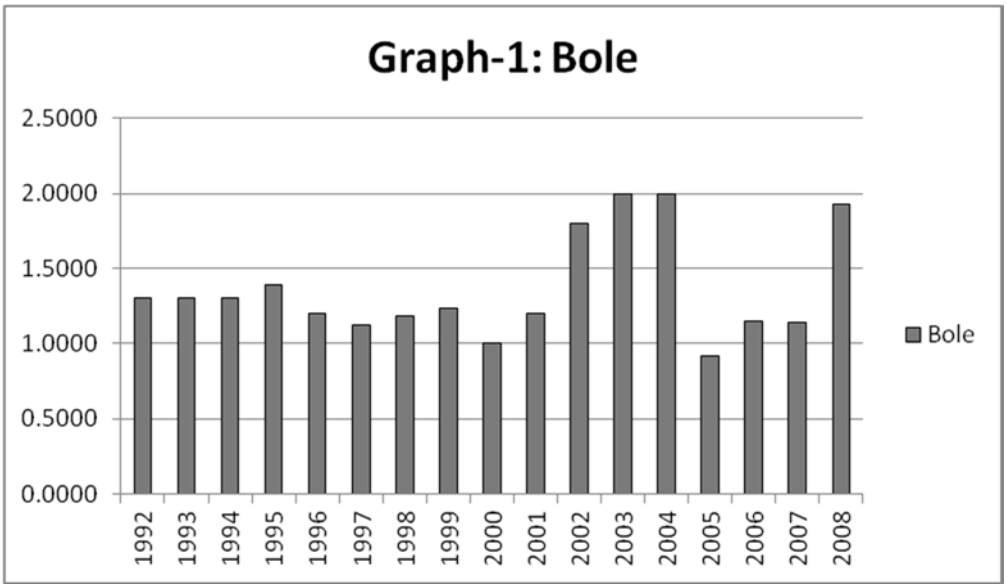
Table-1: Summary Statistics of Maize Yield

Variable	N	Mean	Median	Std Dev	Maximum	Minimum
Bole	17	1.3625	1.2308	0.3475	2.0000	0.9134
Damango	17	1.2744	1.2000	0.5060	2.2898	0.1200
Salaga	17	1.1897	1.2006	0.3821	2.0000	0.4433
Tamale	17	1.1386	1.0000	0.3683	1.9000	0.6000
Yendi	17	1.1678	1.1000	0.2270	1.5785	0.7000

Note: Crop yield was measured in Metric Tons per Hectare (Mt/Ha) in Ghana. Where, 1 hectare = 2.471 acres.

Thus, to avoid these complexities we apply univariate time series methods to achieve simplicity in the model construction. In this paper, crop yield forecasting refers to univariate regional yield forecasts, i.e. forecasting of crop yield (metric tons of crop production per hectare) over large areas. The areas are administrative units called districts, as this is the scale at which most socioeconomic data and crop statistics are available to decision makers.

Table-1presents summary statistics of crop yield for five different districts in Ghana. Univariate time series methods were applied to predict the crop yield (Maize) using seventeen years of data. Average maize yields are more or less similar between districts. However, much variation exists in the maize yield between districts. Even though there are some similar trends observed in the yield plot over time (see, Graphs 1-5), the pattern is not systematic among the districts. As for example, “Damango” district has nine years of downtrend of crop yield that ended in 2003 (see, Graph-2). Similar down trend also exists with other districts that has ended in earlier years and thus makes these patterns non-systematic. To overcome these complex trend movements, we developed time series forecasting models that are applied separately to each district individually to capture the data pattern for that specific region. The following describes the concepts of different time series models briefly, which we have implemented in this research.



Simple Exponential Smoothing:

This technique is based on a series of averaging data in a decreasing (exponential) manner. The weights α is termed the smoothing constant which ranges from 0 to 1. The value of α determines the extent to which the most current observation influences the forecast. The simple exponential smoothing equation is expressed as,

$$L_t = \alpha y_t + (1 - \alpha) L_{t-1} ,$$

where L_t the smoothed value for year t becomes the forecasted value for year t+1.

Table-2: Maize in Bole

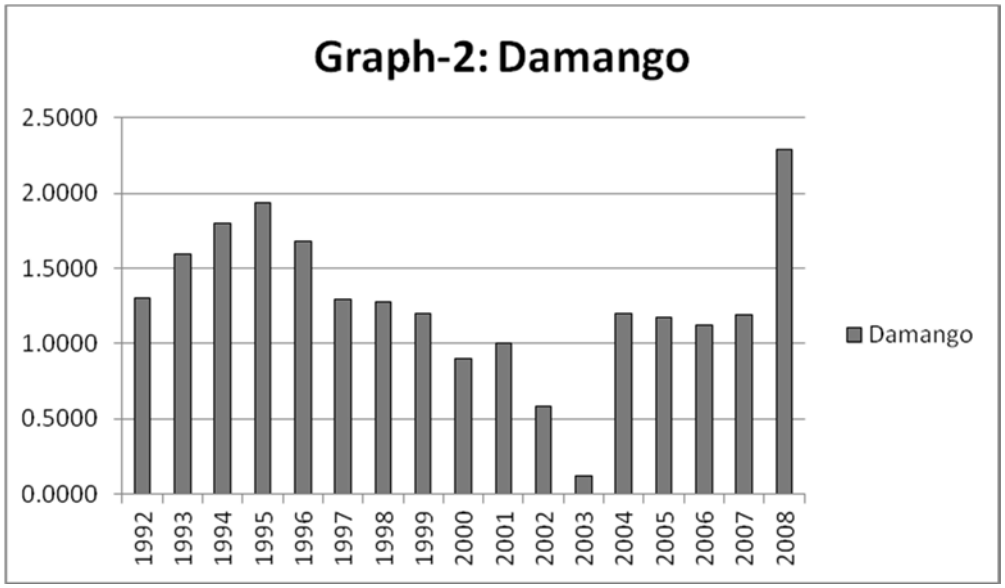
Model	DF	Error Variance (MSE)	AIC	SBC	R-Square
AR(4)	12	0.0787606	-39.12403	-34.95797	0.51
Simple Exponential Smoothing	15	0.1384718	-30.66604	-29.89345	-0.1
Double (Brown) Exponential Smoothing	14	0.1278583	-29.88738	-29.17933	-0.3
Linear (Holt) Exponential Smoothing	13	0.1848473	-23.46989	-22.05379	-0.36
Damped-Trend Linear Exponential Smoothing	13	0.1485906	-27.82719	-25.50943	-0.09

Double Exponential Smoothing (Brown):

A double smoothing technique is used when a series has a trend component. With this technique, each observation in a series is assumed to be consisted of two components, level or smoothing

component and trend component. This controls any trend or nonstationary component that may exist in the data series.

$L_t = \alpha y_t + (1-\alpha)(L_{t-1} + T_{t-1})$ and $T_t = \alpha(L_t - L_{t-1}) + (1-\alpha)T_{t-1}$



Double Exponential Smoothing (Holt):

Holt smoothing technique is different from Brown's technique in a sense that it uses different parameter value for estimating the trend component.

$L_t = \alpha y_t + (1-\alpha)(L_{t-1} + T_{t-1})$ and $T_t = \gamma(L_t - L_{t-1}) + (1-\gamma)T_{t-1}$

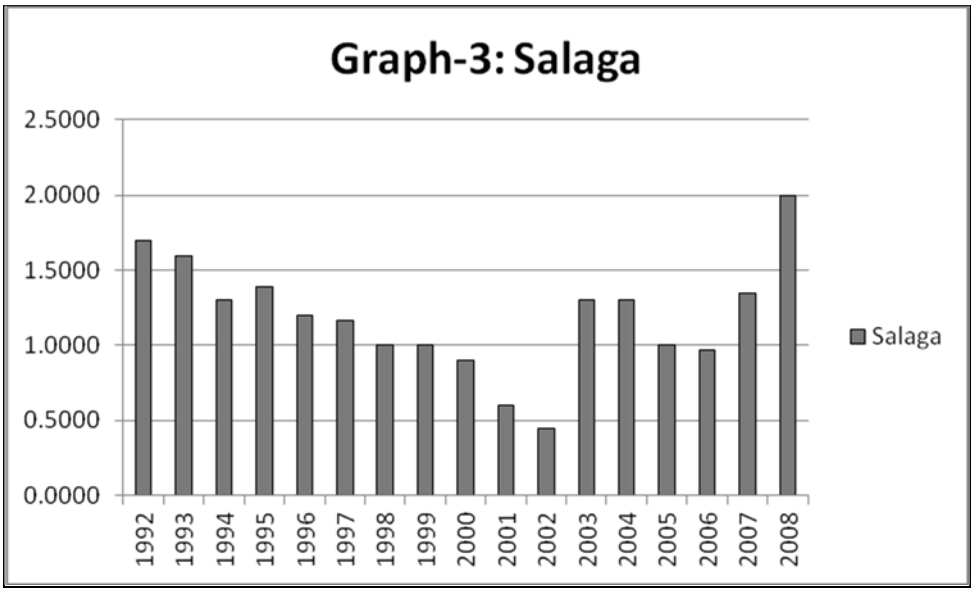
Table-3: Maize in Damango

Model	DF	Error Variance (MSE)	AIC	SBC	R-Square
AR(5)	11	0.1509526	-27.54382	-22.54454	0.556
Simple Exponential Smoothing	15	0.2150581	-23.62217	-22.84958	0.212
Double (Brown) Exponential Smoothing	14	0.250109	-19.82277	-19.11472	0.112
Linear (Holt) Exponential Smoothing	13	0.2428938	-19.37348	-17.95738	0.188
Damped-Trend Linear Exponential Smoothing	13	0.2481439	-19.62217	-17.3044	0.212

Damped-Trend Linear Exponential Smoothing:

This smoothing technique is a variation of Holt smoothing technique that introduces a third parameter value to dampen the trend magnitude to align with a subdued trend data series. This works better with a data series that has weaker trend component.

$L_t = \alpha y_t + (1-\alpha)(L_{t-1} + \varphi T_{t-1})$ and $T_t = \gamma(L_t - L_{t-1}) + (1-\gamma)\varphi T_{t-1}$



Autoregressive Model – AR (P):

Autoregressive (AR) model is a special case of ARMA model of Box-Jenkins (Box and Jenkins, 1976) approach with a stationary data series.

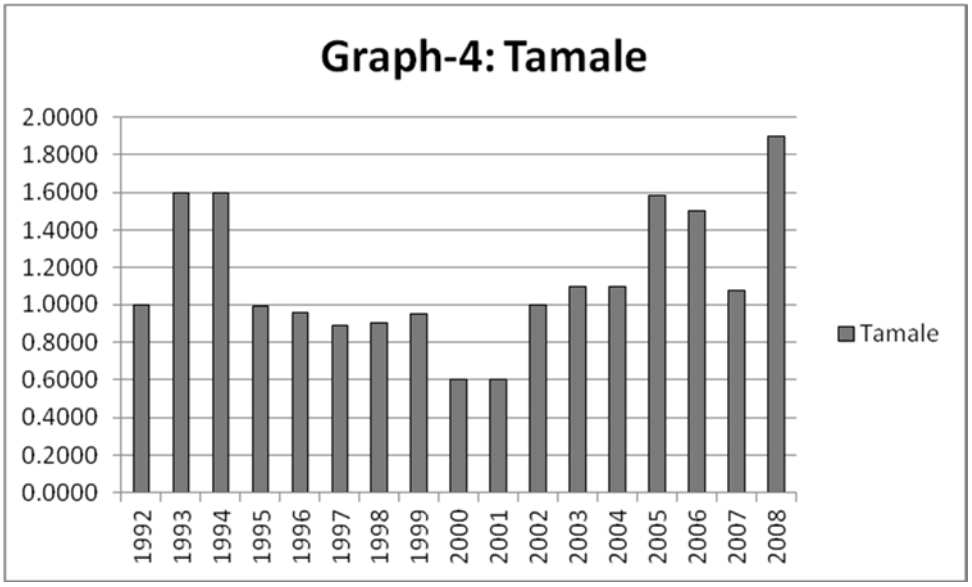
$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + \nu_t$$

Table-4: Maize in Salaga

Model	DF	Error Variance (MSE)	AIC	SBC	R-Square
AR(6)	10	0.0814776	-37.64694	-31.81445	0.624
Simple Exponential Smoothing	15	0.112847	-33.94017	-33.16758	0.178
Double (Brown) Exponential Smoothing	14	0.1339997	-29.18366	-28.47561	-0.04
Linear (Holt) Exponential Smoothing	13	0.1392374	-27.72014	-26.30404	0.002
Damped-Trend Linear Exponential Smoothing	13	0.1302081	-29.94017	-27.6224	0.178

To identify the order of the autoregressive model, we have evaluated the autocorrelation function (ACF) and partial autocorrelation function (PACF) of the yield series using SAS procedure PROC ARIMA (see, SAS/ETS User's Guide, 1993). This allowed the observance of the degree of autocorrelation and the identification of the order of the model that sufficiently described the autocorrelation. After evaluating the ACF and PACF, the models are identified as fourth order to sixth order autoregressive models for various districts and a sixth order model is expressed as:

$\left(1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3 + \phi_4 B^4 + \phi_5 B^5 + \phi_6 B^6\right) \nu_t = y_t$, (see, Box, Jenkins, & Reinsel, 1994) .



Maximum likelihood estimation method is used instead of nonlinear least squares to estimate the parameters of the models. Maximum likelihood estimation is preferable over nonlinear least squares, because maximum likelihood estimation accounts for the determinant of the variance-covariance matrix in its objective function (likelihood function). Further discussion on different estimation methods and the likelihood functions can be found in Choudhury, Hubata, & St. Louis, 1999 and also see SAS/ETS User's Guide, 1993 for the expression of the likelihood functions.

Table-5: Maize in Tamale

Model	DF	Error Variance (MSE)	AIC	SBC	R-Square
AR(6)	10	0.0562229	-43.95401	-38.12152	0.582
Simple Exponential Smoothing	15	0.1277286	-31.95818	-31.18559	0.084
Double (Brown) Exponential Smoothing	14	0.1477733	-27.71603	-27.00798	-0.4
Linear (Holt) Exponential Smoothing	13	0.1255981	-29.26653	-27.85043	-0.16
Damped-Trend Linear Exponential Smoothing	13	0.1456325	-28.14894	-25.83117	0.083

We have used the following model selection criterion:

Akaike Information Criterion: $AIC = n \ln \frac{SSE}{n} + 2k$,

Schwartz's Bayesian Criterion: $SBC = n \ln \frac{SSE}{n} + k \ln(n)$,

and R-Square = $1 - \frac{SSE}{SST}$. Note that in this construct of R^2 the value of R^2 can be negative when the fitted model's performance is very poor. This means that the total squared deviation of predicted yield from actual yield is larger than the total squared deviation of average yield from actual yield.

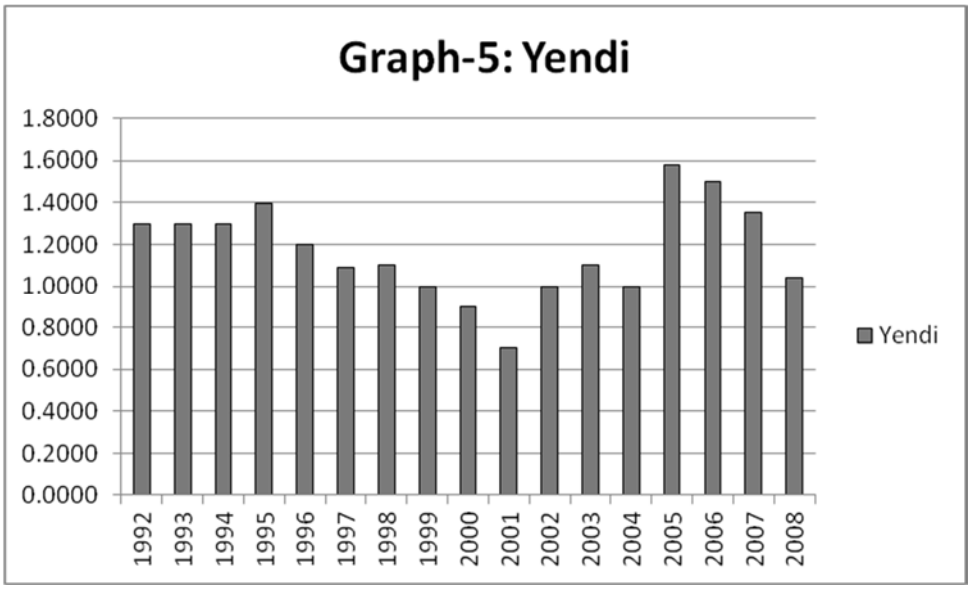
EMPIRICAL RESULTS

Larger standard deviation 0.5060 of Maize yield in "Damango" district with the highest "maximum" yield of 2.2898 and lowest "minimum" yield of 0.1200 (see Table-1) does indicate much fluctuation in the Maize yields among the districts and thus introduces a challenge in model building strategy. Average "maximum" yield is about 2.0 Mt/Ht among these five districts, whereas the "minimum" yield varies quite a bit with a range of 0.9134 Mt/Ht to 0.1200 Mt/Ht. There appears to be a declining trend in Maize yield till 2001/2002, followed by an increasing trend in yield for a period of six/seven years (see, Graphs 1-5). This may be one of the reason why trend adjusted forecasting (or smoothing) technique is performing so poorly and thus producing negative R^2, which is essentially zero. Since, theoretically coefficient of determination ranges from zero to one and cannot be negative. Thus, it appears that there are possibly two opposite crop yield trends which create a cycle that split up around the year 2001/2002. It is possible that this may be due to weather cycle occurrence or management intervention or some other unobservable phenomena of similar nature. This cycle may be country specific and may also be region specific and therefore, needs to be explored further in the future research.

In a similar context, there are also visible differences in declining trend segment ending in a different time period for different districts and thus exhibiting differences in external factors' influence on the crop yield differently. This suggests that due to some unobservable factor(s) crop yield may differ in different time periods for different districts. Thus, the idea of this exploratory analysis is to obtain a best fit forecasting model of crop yield such that the association effect of unobserved external factors with crop yield is best reflected through models' performance criterion. The following results address our research studies of building the forecasting model of crop yields for different districts.

Among all different time series models estimated, ARMA models performed best with higher coefficient of determinations for all five districts considered in this paper. AR (6) model fitted for district "Salaga" has the highest R^2=0.624 that accounts for 62.40% variation (see, Table-3) in the Maize yield. The model that produced the lowest R^2, among these five districts is "Yendi". District "Yendi" fitted a model that accounts for 48.70 % variation (see, Table-5) in the crop yield. In addition, ARMA models also performed best when considering other performance criterion, such as, MSE or Akaike Information Criterion (AIC). Therefore, our research results show that

ARMA model provides better estimate of crop yield using historical data at a district level compared to other models considered in this paper.



It appears that in addition to the plant characteristics, external factors may also affect the crop yield differently given that which time period they are planted. Specifically, we observe that there exists a crop yield cycle in most of our data sets, which starts with the downtrend that ended around 2001/2002 and then an uptrend for next several years that creates a crop yield cycle. In general, any type of time series data has a cyclical component whether it is visible or subdued. A number of possible explanations can be explored for this time dependent yield cycle. However, considering that most of the time series has some serial correlation properties inherent in them direct comparison may be complicated and difficult to separate.

CONCLUSION

This paper makes a number of significant contributions to the literature. It provides additional evidence of crop yield cycle component of a time series in most of the districts. In addition, it also suggests evidence of unobserved external factors' effect on crop yield that creates the crop yield cycle. However, any associations of crop yield that may exist with the unobserved external factors' are not explored in this study. These results while important are not unexpected given the dynamic changes that come from external factors, such as, weather (rainfall, temperatures, etc.), land management (that include re-division of districts), pests and diseases. The unexpected finding is the initial continuous decline of crop yield that went on for several years in most of the districts without any management intervention.

Table-6: Maize in Yendi					
Model	DF	Error Variance (MSE)	AIC	SBC	R-Square
AR(4)	12	0.0345774	-53.11867	-48.9526	0.487
Simple Exponential Smoothing	15	0.0455613	-48.45175	-47.67916	0.152
Double (Brown) Exponential Smoothing	14	0.0624674	-40.63155	-39.9235	-0.12
Linear (Holt) Exponential Smoothing	13	0.0525918	-42.32445	-40.90835	0.035
Damped-Trend Linear Exponential Smoothing	13	0.0525708	-44.45175	-42.13399	0.152

Considering crop yield trend and crop yield cycle separately from other factors (external or internal) and purely from the historical point of view, illustrates how policy makers can benefit from using the results of this study. It is also well known that most of the time series has an inherent cycle component that may or may not be significant. However, understanding the mechanism of up-cycle and down-cycle with crop yield will provide an advantageous position to the policy makers to prepare an appropriate policy design for yield management.

Therefore, a successful operation of an Area-Yield Index insurance policy to work the crop grown in the Insured Unit (District) needs to be relatively homogeneous in terms of the varieties grown by farmers, sowing dates, crop husbandry practices and input utilization and finally the average yields of the crop obtained by the farmers in the defined unit. To date no work has been conducted on individual crop-cut yields to assess the degree of variability in crop yields obtained by farmers in the same district. Additional research development is needed, particularly with regard to the linkage between these factors and crop yield dynamics. To determine the length of downtrend or uptrend and therefore the total cycle of crop yield, future research could examine these phenomena over different periods of time.

The ARMA models, which are univariate models that use primarily autocorrelations from its past, proved more robust time-series models than the smoothing technique models for predicting crop yield in this study. This is consistent with the findings of Pal et al., 2007. They found that an ARMA model for forecasting Milk production resulted in much better estimates than the other time series approaches considered. The ARMA methodology avoided the problem of highly variable crop yields within the district over time, which has led to low performance on the prediction of crop yield by averaging or smoothing models. This predictive power of ARMA models does not depend on whether and how long the crop yield cycle persists. These findings are consistent with the objective that an efficient prediction modeling process is very much interrelated with the yield data itself. Therefore, the results of this study indicate that the performance of a prediction model is dependent on the dynamic nature of the crop yield data and this may be region

specific. Thus, the districts with wider yield spreads may like to use different time series models than those districts with more homogenized yield.

REFERENCES

Armstrong, J.S. (Ed), 2001. *Principles of forecasting, a handbook for researchers and practitioners*. Kluwer Acad. Press, Boston, Doordrecht, London.

Badmus, M.A., and Ariyo, O.S. (2011) 'Forecasting Cultivated Areas and Production of Maize in Nigeria Using ARIMA model', *Asian Journal of Agricultural Sciences* 3(3), 171- 176.

Barnett, B. J., Barrett, C. B. and Skees, J. R. (2008) 'Poverty traps and index-based risk transfer products', *World Development* 36(10), 1766-1785.

Bogetic, Y., Bussolo, M., Ye, X., Medvedev, D., Wodon, Q., and Boakye, D. (2007) 'Ghana's growth story: How to accelerate growth and achieve MDGs', *Background paper for Ghana CEM.*

Boken, V.K., 2000. Forecasting spring wheat yield using time series analysis: a case study for the Canadian Prairies, *Agronomy Journal*, 92(6),1047-1053.

Box,G. E. P.and Jenkins, G.M. (1976) *Time Series Analysis: Forecasting and Control*, San Francisco: Holden-Day, Inc.

Box, G.E.P., Jenkins G.M., and Reinsel G.C. (1994) *Time Series Analysis: Forecasting and Control*, Englewood Cliffs, NJ: Prentice-Hall.

Carbone, G. J., Kiechle, W., Locke, C., Mearns, L. O., McDaniel, L. and Downton, M. W. (2003) 'Response of soybean and sorghum to varying spatial scales of climate change scenarios in the southeastern United States', *Climatic Change* 60(1 - 2), 73–98.

Carter, M. R., Galarza, F., and Boucher, S. (2007) 'Underwriting area-based yield insurance to crowd-in credit supply and demand', *Savings and Development* 31(3), 335-362.

Choudhury, A., Hubata, R. and St. Louis, R. (1999) 'Understanding Time-Series Regression Estimators', *The American Statistician*, 53(4), 342-348.

Deng, X., Barnett, B. J., Hoogenboom, G., Yu, Y., and Garcia, A. (2006) Evaluating the efficiency of crop index insurance products, Paper presented at the *Southern Agricultural Economics Association annual meetings*, Orlanda, Florida.

Easterling, W. E., Weiss, A., Hays, C. J. and Mearns, L. O. (1998) 'Spatial scales of climate information for simulating wheat and maize productivity: the case of the US Great Plains', *Agricultural and Forest Meteorology* 90(1-2), 51–63.

Etwire, P.M., Al-Hassan, R.M., Kuwornu, J.K.M. and Osei-Owusu, Y. (2013) 'Application of Livelihood Vulnerability Index in Assessing Vulnerability to Climate Change and Variability in Northern Ghana', *Journal of Environment and Earth Science* 3(2),157-170.

Fodor, N. and Kovacs, G. J. (2005) 'Sensitivity of crop models to the inaccuracy of meteorological observations', *Physics and Chemistry of the Earth*, Parts A/B/C 30(1-3), 53–57.

Gerretsadikan, A and Sharma, M.K. (2011) 'Modeling and forecasting of rainfall data of mekele for Tigray region (Ethiopia)', *Statistics and Applications* 9(1-2), 31-53.

Hansen, J.W., Challinor, A., Ines, A., Wheeler T. and Moron, V. (2006) 'Translating climate forecasts into agricultural terms: advances and challenges', *Climate research* 33,27-41.

Hansen, J. W. and Jones, J. W. (2000) 'Scaling-up crop models for climate variability applications', *Agricultural Systems* 65(1), 43–72.

Hoogenboom, G. (2000) 'Contribution of agrometeorology to the simulation of crop production and its applications, *Agricultural and Forest Meteorology,* 103(1-2), 137–157.

Launay, M. and Guérif, M. (2003) 'Ability for a model to predict crop production variability at the regional scale: an evaluation for sugar beet', *Agronomie* 23, 135–146.

Lawless, C. and M.A. Semenov, 2005. Assessing lead-time for predicting wheat growth using a crop simulation model. *Agricultural and forest meteorology* 135(1-4),302-313.

Makridakis, S., Wheelwright S.C. and Hyndman, R.J. (1998) *Forecasting, Methods and Applications*, New York: John Wiley & sons, Inc.

Mearns, L. O., Mavromatis, T., Tsvetsinskaya, E., Hays, C. and Easterling, W. (1999) 'Comparative responses of EPIC and CERES crop models to high and low spatial resolution climate change scenarios', *Journal of Geophysical Research-Atmospheres* 104(D6), 6623–6646.

Mearns, L. O., Easterling, W., Hays, C. and Marx, D. (2001) 'Comparison of agricultural impacts of climate change calculated from high and low resolution climate change scenarios: Part I. the uncertainty due to spatial scale', *Climatic Change* 51(2), 131–172.

Najeeb, I., Khuda, B., Asif, M., and Abid, S.A. (2005) 'Use of ARIMA Model for Forecasting Wheat Area and Production in Pakistan', *Journal of Agricultural and Social Sciences* 1(2), 120- 122.

Nonhebel, S. (1994) 'The effects of use of average instead of daily weather data in crop growth simulation models', *Agricultural Systems* 44(4), 377–396.

Pachepsky, Y. and Acock, B. (1998) 'Stochastic imaging of soil parameters to assess variability and uncertainty of crop yield estimates', *Geoderma* 85(2-3), 213–229.

Pal, S., Ramasubramanian, V. and Mehta, S.C. (2007) 'Statistical models for forecasting milk production in India', *Journal of Indian Society of Agricultural Statistics* 61(2), 80–83.

Russell, G. and Gardingen, P.R.v. (1997) 'Problems with using models to predict regional crop production', in Gardingen, P.R. v., Foody, M. and Curran P. (eds), Scaling up: from cell to landscape, Cambridge University Press, Cambridge, pp. 273–294.

Semenov, M. A. and Porter, J. R. (1995) 'Non-linearity in climate change impact assessments', *Journal of Biogeography* 22(4-5), 597–600.

Skees,J.R., Black, J.R. and Barnett, B.J. (1997) 'Designing and rating an area yield crop insurance contract', *American Journal of Agricultural Economics* 79, 430-438.

Skees, J. R. (2008) 'Innovations in index insurance for the poor in lower income countries', *Agricultural and Resource Economics Review* 37(1),1-15.

Soltani, A., Meinke, H. and de Voil, P. (2004) 'Assessing linear interpolation to generate daily radiation and temperature data for use in crop simulations', *European Journal of Agronomy* 1(2), 133–148.

Stutley, C. (2010) Innovative Insurance Products for the Adaptation to Climate Change Project Ghana, Innovative Insurance Products for the Adaptation to Climate Change Project Ghana.

Suleman, N. and Sarpong, S. (2012) 'Forecasting Milled Rice Production in Ghana Using Box-Jenkins Approach', *International Journal of Agricultural Management and Development*, 2(2), 79-84.

THE CPI MARKET BASKET: A REVIEW OF ECONOMIC AND MARKETING VALIDITY ISSUES

James J. Csipak, SUNY College at Plattsburgh
Cataldo Zuccaro, University of Quebec at Montreal

ABSTRACT

This article reviews economic and marketing validity issues of the Consumer Price Index (CPI) and the CPI Food at Home market basket as computed and reported by the Bureau of Labor Statistics' (BLS). It identifies a major controversy that has plagued the CPI for more than 70 years as well as addresses economic as well as marketing issues which explain the CPI's shortcomings. Specifically, it reviews the CPI literature for both food and non food components and reviews the major issues surrounding the CPI's computation methodology used by the BLS, and puts forth recent recommendations that have greatly helped reduce many of the CPI's shortcomings.

Keywords: Consumer Price Index, Bureau of Labor Statistics.

INTRODUCTION

The original impetus for the present article came from our School of Business and Economics' Marketing and Entrepreneurship Department (within the State University of New York College at Plattsburgh) acceptance, this in summer 2011, to take over the "Food Market Basket Data collection project". This project had been previously accomplished by a now defunct on-campus federally funded agency. This agency had been in charge of collecting, measuring, and reporting food price fluctuations by surveying three (3) conventional supermarkets and one Wal-Mart Supercenter, all located in the city of Plattsburgh, a rural setting (population of 22,000 people) in upstate New York, near the U.S./Canada border. The survey instrument utilized was composed of forty-one (41) food items. In September 2011, a quick perusal of the Consumer Price Index (CPI hereafter) food at home literature informed us that, within the last 20 years, major changes had occurred within food at home purchases by U.S. consumers. These changes, as reported by MacDonald (1995), were: 1. Shifts in consumer behavior such as in the case of decreased purchases in the food-at-home category because of increased purchases at restaurants; 2. Shifts in types of food purchases such as the purchasing of more "fresh" fruits and vegetables and less meat products; 3. Shifts in the Amount of new food products introduced in Supermarkets (for example, the number of new products introduced in Supermarkets increased from 5,400 in 1984 to 12,300 in 1992) and 4. Shifts in the amount and types of new retail outlets that sell food as in the case of a growing share of food sales occurring outside conventional supermarkets such as at drug stores, at warehouse club stores, at mass merchandisers (or general discount retailers), and at convenience stores as well. Because we were informed that our inherited 41-food item survey instrument dated back to 1978, we suspected that these issues and shifts had not been accounted for. Our team agreed that an assessment of our survey instrument's validity was in order. Our initial quick perusal of the CPI literature had equally revealed that there were a number of very important validity issues

as to how the Bureau of Labor Statistics (BLS hereafter) computed the CPI that remained unresolved to this day. We decided that we would conduct an exhaustive literature review of both the CPI food at home category as well as the CPI's other goods and services since CPI validity issues would concern all products surveyed by the BLS, the federal government agency responsible for computing and publishing the CPI on a monthly basis. As we are marketing scholars and had neither previous knowledge nor experience with the CPI, we believed this effort would help us, first, to best understand the "benchmark" of price fluctuation indexes in the U.S. and, second, help us make improvements to our survey instrument.

WHY THE CPI MATTERS

As stated by Schultze and Mackie (2002) "the Consumer Price Index (CPI) is one of the most widely used statistics in the United States. As a measure of inflation it is a key economic indicator. It serves as a guide for the Federal Reserve Board's monetary policy and is an essential tool in calculating changes in the nation's output and living standards. It is used to determine annual cost-of-living allowances for social security retirees and other recipients of federal payments, to index the federal income tax system for inflation, and as the yardstick for U.S. Treasury inflation-indexed bonds." Invariably, as suggested by Boskin et Al. (1998) the CPI impacts the U.S. national budget and the national debt as well.

A DESCRIPTION OF THE CPI

Essentially, the CPI is a measure of the average change in prices paid by urban consumers for a fixed market basket of goods and services including food" (MacDonald, 1995).
According to Wahl (1982) the CPI is "simply a fixed-weight index for measuring changes in consumer prices between a base period and a subsequent period, the weights being established by the typical expenditures of all consumers in the base period". The U.S. Bureau of Labor Statistics first published an index of consumer prices for food at home in 1903 (Reinsdorf and Triplett, 2004). In 1919, the CPI was expanded to include cloth and apparel. Equally occurring in 1919, the CPI started using weights which were computed from surveys of consumer purchasing patterns, this so as to distinguish the level of expenditures or relative importance of each product category included in the CPI. According to Reinsdorf and Triplett (2004), the history of economists' analysis of BLS prices indexes began with the assistance of Irving Fisher and Wesley Mitchell with the food index improvements of 1914. It is Mitchell who, in 1915, drew attention to the fact that the CPI incorrectly used Wholesale Price data (being more easily available) instead of Consumer Price data (Goldberg and Moye, 1985). Hence, at that point in time, the CPI was not accurately measuring price inflation at the consumption level, although the procedure of using wholesale price data may be excused since methods to survey the retail sector were far from being well developed.

As stated on the BLS website (WWW.BLS.GOV/CPI), the CPI is a monthly measure of the average change of the prices of eleven goods and services categories (see Table 1 at the end of the paper) and is published for two population groups: (1) the CPI for Urban Wage Earners and Clerical Workers (CPI-W) which comprise approximately 29 percent of the total population and

(2) the CPI for All Urban Consumers (CPI-U) and the Chained CPI for All Urban Consumers (C-CPI-U) which covers approximately 88 percent of the population. Every month, the BLS collects prices in 87 urban areas across the country from about 4,000 housing units who keep records of what products they have purchased and what prices they have paid. The BLS then verifies the prices paid for these products by trained representatives who either call or visit approximately 26,000 retail establishments such as department stores, supermarkets, hospitals, filling stations, and other types of stores and service establishments. In calculating the CPI, price changes for the various items in each of the 87 locations are averaged together with weights, which represent their importance in the appropriate population group. The weights of all goods and services equal 100. The weights of each good and/or service are subject to change over time. Food, for example, may at one point in time have a weight of 14.208 out of the total of 100. But if consumers find themselves spending more for energy because of increases in the cost of fuel or electricity, their food expenditures may change. The CPI measures price changes from a designed reference date. For the CPI-U and CPI-W the reference base is 1982-84 which equals 100. In the case of the C-CPI-U the reference base is December 1999 which equals 100. An increase of 16.5 percent from the reference base, for example, will be shown as 116.500.

The major challenge faced by the BLS in computing the CPI is to accurately compute a price index that is "representative" of "true" price changes of products purchased by consumers in U.S. Markets. In view of products, it must first chose to survey the prices of a representative sample of products being purchased, taking into account "new products" entering the market which may affect the prices of existing products. It must equally take into consideration the "quality improvements" of some products which may also affect the prices of existing products. It also needs to select a representative sample of retail outlets that includes new types of outlet formats which may sell products at a different price level and affect the prices at which products are sold at conventional outlets. For example, in view of choosing a representative sample of existing retail outlets as well as incorporating new types of outlets, the BLS uses a survey called the telephone point-of-purchase (TPOPS). Outlets are sampled from the TPOPS frame in proportion to their estimated sales within each of the goods and services category.

In essence, if the BLS desires that the CPI accurately measure and report price changes in the goods and services it surveys, it needs, for the least, to insure the following four aspects: first, that the "level of expenditures" per good and service category surveyed is representative of consumer purchasing behavior; second, that "new products" that enter the market be represented in the samples of goods and services surveyed; third, that "quality changes" in products be represented and accounted for in the samples of goods and services surveyed; and finally, that "new retail shopping outlets" be included in the outlet sample so that their impact on prices paid by consumers be accounted for. Hence, updating the CPI to correctly account all of these four market changes begin occurring, this on a non-delay basis is at the heart of the CPI's accuracy and, as we shall see, what characterizes the CPI's historical development.

CHALLENGES FACED BY THE BLS IN COMPUTING THE CPI

In their final report, the 1995 Boskin Commission clearly acknowledged the multitude of major challenges faced by the BLS in computing the CPI by stating the following: "Hence, the

very first point the CPI Commission made in its report was that inflation was inherently difficult to measure in a complex dynamic market economy" (Boskin et Al., 1998). Their answer to the question why changes in the cost of living are so hard to measure was most revealing of these challenges: "How to obtain information on who is buying what, where, when, and how in an economy, and then to aggregate it into one or a few measures of price changes raises a host of complex analytical and practical problems" (Boskin et Al., 1998). To illustrate the momentous challenges faced by the BLS, they put forth the following 4 points: First, they noted that there were literally millions of goods and services available in a modern economy and that for example, a single supermarket could contain some 30,000 differently priced items and that a Wal-Mart store could offer over 40,000 priced items; Second, that new products were being constantly introduced while existing ones were either being improved or disappeared; Third, that relative prices of different goods and services changed frequently in response to technological innovations and other factors affecting costs and quality which in turn led consumers to change their buying patterns; Finally, that as the U.S. had become richer, demand had increasingly shifted from goods to services, and, as well, to characteristics of goods and services such as enhanced quality, more variety, and greater convenience. They concluded: "But all these factors, plus others, means a larger fraction of what is produced and consumed in an economy is harder to measure than decades ago when a large fraction of economic activity consisted of a smaller number of easier to measure items such as hammers and potatoes". Interestingly, the 1995 Boskin Commission, in its final report, acknowledged the many challenges faced by the BLS, by stating: "The dimensionality of this task is difficult to convey, and we would not wish our recommendations for improvements to detract from our admiration for, and support of, the basic program as it has evolved and improved over the years."

THE CPI'S FOUR MAJOR PROFESSIONAL REVIEWS

A review of the CPI literature by Reinsdorf and Triplett (2004) suggests that between the 1930's and 2002, four major professional reviews of the CPI took place, each constituting a milestone in the history of the CPI. According to them, although a professional review of the CPI took place in 1933-34 by an Advisory Committee appointed by the American Statistical Association, the 1944 Mitchell Committee corresponds to the CPI's first professional review. This, they contend, was because the 1944 Mitchell Committee's recommendations were not only practical but equally conceptual with more in-depth discussions as well as larger in scope. The second major professional review was conducted in 1961 by the Stigler Committee which was appointed by the U.S. Bureau of the Budget. The third major professional review was made by the 1995 Boskin Commission which was appointed by the Senate Finance Committee. Finally, the fourth and latest major professional review was conducted by the 1999 Committee on National Statistics of the National Academy of Sciences Panel (CNSTAT panel hereafter) which was funded by the BLS and, was largely a response to the 1995 Boskin Commission review. In the following paragraphs, we shall summarize each review's impact on the CPI.

THE 1944 MITCHELL COMMITTEE: THE CPI UNDERESTIMATES INFLATION

The first professional review of the CPI according to Reinsdorf and Triplett (2004) took place under the auspices of the 1944 Mitchell Committee. Interestingly, as reported by Reinsdorf and Triplett (2004), initially, the most vocal CPI critics were the authors of the 1944 Meany-Thomas report, who held that the BLS "understated" price inflation in computing the CPI. This occurred during the Second World War when, because of necessary shortages/rationing in support of the War effort, prices were under inflationary strains. The U.S. government had to intervene by setting price controls as well as price subsidies to maintain the affordability of many goods. It is at that time that the CPI controversy gained national coverage, this especially with the 1944 Meany-Thomas report that vehemently disputed the CPI's accuracy in reporting price inflation. Examining the period from January 1941 to December 1943, George Meany and R.J. Thomas, respectively from the American Federation of Labor (AFL) and the Congress of Industrial Organizations (CIO), calculated that food price inflation had been 74.2% during that period as compared with the BLS's CPI reported rate of 40.2%, which, in effect, according to them, "understated" food price inflation by 34%. In terms of all goods and services measured by the CPI, the BLS reported a 23.4% rise in the cost-of-living for that same period of time whereas the 1944 Meany-Thomas calculated a 43.5% rise in prices.

It is at that point in time that the 1944 Mitchell Committee was formed and concluded that the effects of the unaccounted sources of biases within the CPI identified by the 1944 Meany-Thomas report were, either much smaller than claimed, or, that they were absent. As stated by Reinsdorf and Triplett (2004), the BLS vigorously disputed the Meany-Thomas report's contention that the cost-of-living index underestimated inflation by almost half, a position that was largely supported by the outside experts on the 1944 Mitchell Committee. Most interestingly, the 1944 Mitchell Committee recommended that the BLS change the name of its CPI index. The new name adopted by the BLS in September 1945 became "The Consumers' Price Index for moderate Income Families in large cities". Importantly, the 1944 Mitchell Committee, in response to the Meany-Thomas report estimates, produced their own estimates of the probable size of the CPI error. This is something that would equally be done later by the 1995 Boskin Commission, with the difference that they would claim that inflation was "overstated" rather than understated, as was claimed by CPI critics during the Second World War (Reinsdorf and Triplett, 2004). The 1944 Mitchell Committee concluded that, for the January 1941 to December 1943 period, the combined effects for all the unaccounted sources of bias by the CPI, as stated in the Meany and Thomas report, might have been no more than 3 to 5 percentage points above the CPI rate reported by the BLS.

The Meany-Thomas report is important because it clearly identified and discussed a number of unaccounted biases in how the BLS measured the CPI which, according to them, precluded it from accurately measuring and reporting price inflation. Most importantly, these biases have been re-addressed numerous times over the last 70 years and remain relevant to this day (although the CPI is now judged by its critics, this especially since the late 1970's, to overestimate price inflation). As stated by Reinsdorf and Triplett (2004), the 1944 Meany-Thomas report identified the following five "unaccounted for market conditions" by the CPI which in their

case, they claimed, contributed in having the CPI "underestimate" inflation : First, they contended that, during the January 1941 to December 1943 period, consumers were often forced to "substitute" more expensive varieties of goods for ones that had disappeared from the marketplace because of wartime shortages or "product line upgrading"; second, they alleged that consumers were often forced into more expensive dwellings because of shortages of affordable housing; third, lower-quality varieties of products often replaced higher-quality ones which occurred when manufacturers relabeled a lower grade item as a higher grade one (they gave the example of the deterioration in the quality of shoes manufactured during the wartime effort); the fourth and fifth unaccounted market conditions by the CPI argued by Meany and Thomas, were about "forced lifestyle changes" that incurred additional costs for consumers such as increased consumption of restaurant meals due to "meat rationing" as well as the entrance of women into the labor force and, extra costs incurred by consumers from migrating between cities to fill wartime jobs.

Historically, the CPI was referred to as a "cost-of-living index" (COLI hereafter). An important part of the 1944 Mitchell Committees' response to the 1944 Meany-Thomas report was a clarification of the conceptual goal of the BSL's cost-of-living index, as the CPI was often referred to at the time (Reinsdorf and Triplett, 2004). At that point in time, the term "cost-of-living index" was interpreted in at least three different ways: First, as a price index that holds constant the cost of living (which corresponds to the current standard interpretation); Second, as a fixed basket index that covered a family's entire budget; and Third, as the cost of attaining the standard of living deemed appropriate compared to the cost of a possibly lower standard of living in some previous period. Most importantly, the 1944 Mitchell Committee's view was that the CPI or cost-of-living index ought to measure only the influence of prices on the cost of living, not the influence of other factors such as those underlined by the 1944 Meany-Thomas Report. Hence, the substitutions that consumers experienced such as "forced uptrading", increased dwelling rents, diminishing quality of goods, and lifestyle changes were to be considered outside of the CPI's domain or realm. It is to clarify that nonprice influences on welfare were out of the CPI's scope as well as to avoid confusing the CPI with an index that measured few changes in the standard of living that the 1944 Mitchell Committee recommended that the BLS change the CPI's name to that of "Consumer Price Index for Moderate Income Families in Large Cities" (Reinsdorf and Triplett, 2004). In fact, at that point in time, the CPI was most representative of a straightforward "cost-of-goods index" (COGI hereafter) than of a COLI.

THE 1961 STIGLER COMMITTEE: THE CPI OVERESTIMATES INFLATION

The second professional review of the CPI according to Reinsdorf and Triplett (2004) took place under the auspices of the 1961 Stigler Committee. As reported by them, in 1957, the Joint Economic Committee conducted an investigation of employment, growth, and price levels. In subsequent hearings, the need for reliable price statistics emerged as a minor theme. A paper by economist Kenneth Arrow, who argued for a COLI objective for the CPI because of the importance of commodity substitution behavior by consumers, caught considerable attention. Subsequently, the U.S. Bureau of the Budget contracted with the National Bureau of Economic Research which appointed, in 1961, a Price Statistics Review Committee chaired by George Stigler. Two major positions taken by the 1961 Stigler Committee summarize its contribution to improving the CPI's

accuracy: first, the Committee agreed with the consensus by participants of the Second World War era who recommended that the CPI ought, in principle, to reflect the effects of substitution as exercised by consumers in buying cheaper goods/brands when these were available. According to Reinsdorf and Triplett (2004), the 1961 Stigler Committee went beyond the 1944 Mitchell Committee in stating unequivocally that the measurement concept for the CPI ought to be the cost of staying on an indifference curve. Second, the Committee was first to discuss a rather new market phenomenon occurring at the time, that of "voluntary substitution", this in contrast with the "forced uptrading" that had characterized Second World War consumer markets. In this, the Committee recognized that the CPI, in using fixed weights for each product category it surveyed for consumer expenditures, was not taking into account the effects of substitution when consumers would purchase alternatives goods or services that would save them money. More importantly, the CPI, in using these fixed weights, could not accurately measure price changes, and in contrast with the CPI's underestimating price inflation during the Second World War, it would now "overestimate" prices. Because the BLS in computing the CPI used a Laspeyres index corresponding to an arithmetic mean that could not account for substitution, the 1961 Stigler Committee recommended that the BLS periodically estimate price changes by using a Paasche index version of the CPI to gage the size of the bias from substitution never accurately accounted for by the CPI. The Paasche index uses a geometric mean that equals price elasticity to 1 instead of 0, as in the case of the Laspeyres index, and assumes the existence of substitution. Hence, when consumers would substitute lower priced items for higher priced items, this effect could be accounted for and consequently price changes would be reported more accurately. The 1961 Stigler Committee not only addressed the unaccounted substitution bias by the CPI but equally the CPI's unaccounted biases of changes in the quality of products, the treatment of consumer durables, and the price effects of new products. A major recommendation put forth by the 1961 Stigler Committee was that the BLS use probability sampling in view of how it chose samples of products and outlets to survey price changes. This procedure, the Committee contended, would be the only way to guard against biases due to an unrepresentative selection in the variety of products as well as retail store outlets. According to Reinsdorf and Triplett (2004), of all the 1961 Stigler Committee's recommendations, this was the one that would have the most important effect in helping improve the CPI's accuracy in measuring price changes.

Another major recommendation by the 1961 Stigler Committee was that the BLS should re-orient the CPI towards a cost-of-living index (COLI) rather than simply reporting price changes as in the case of the "cost-of-goods index" (COGI). This recommendation, if adopted, would gradually move the CPI closer to becoming a "welfare or "constant utility" utility index. As reported by Reinsdorf and Triplett (2004), the BLS's initial reaction to this recommendation was quite negative. Their rationale at that point in time was based on both the lack of research showing how to estimate a COLI and, on the BLS's doubts about the suitability of the COLI for the purpose or objective of the CPI. Interestingly, the BLS would eventually reverse itself and adopt the COLI as a conceptual framework but this would occur three years after the 1995 Boskin Commission, hence, 37 years later.

An article by Janet Norwood (1964) is indicative of how a BLS "supporter" (as opposed to a BLS "critic") perceived most of the recommendations put forth by the 1961 Stigler Committee. Acknowledging that the CPI was being criticized for overstating both the cost of living and

inflation, Norwood (1964), who was Commissioner of Labor Statistics at the time, stated that the CPI was a good measure of the changes in the purchasing power of the average family represented in the CPI index. She equally stated: "The CPI is based on a fixed market basket. That is, the weights of the mix of goods and services purchased during the base period are held constant from year to year until a major revision occurs. We keep the market basket constant deliberately because we want to keep fixed the living standard represented by that market basket. Our purpose, to the extent possible, is to isolate price changes from other changes which may occur in living standards". Interestingly, Norwood (1964) seemed pleased to report that the BLS had recently began to utilize a new consumer expenditure survey program that used the Census Bureau data collection of consumer expenditures as a basis for revising the CPI weights. Hence, her position vis-à-vis the 1961 Stigler Committee suggests that she, like the BLS, supported a COGI rather than a COLI as the CPI's objective.

THE 1995 BOSKIN COMMISSION: THE CPI OVERESTIMATES INFLATION

The third professional review of the CPI according to Reinsdorf and Triplett (2004) took place under the auspices of the 1995 Boskin Commission. It is interesting to note that, as reported by Wahl (1982), from 1965 till February 1982, hence a period of a little more than 17 years, the CPI had not registered one single monthly decline. In a 1995 article, MacDonald estimated that the CPI-U food at home, the nation's principal indicator of changes in retail food prices, "overestimated" inflation of food prices between 1 to 1.9 percentage points per year, this beginning around 1978.

According to Reinsdorf and Triplett (2004), it is after a remark on upward bias in the CPI in a testimony by Federal Reserve Chairman Alan Greenspan that the Senate Finance Committee appointed an Advisory Commission to study the CPI. That commission became known as the Boskin Commission, after its chair, Michael Boskin. Most interestingly, when the Boskin Commission was appointed in 1995, hence thirty-four years after the 1961 Stigler Committee, the BLS had still not implemented most of the 1961 Stigler Committee's major recommendations aimed at bringing the CPI into closer alignment with a COLI (Reinsdorf and Triplett, 2004). For example, the weights used to distinguish relative product category expenditures were not updated frequently enough; new goods that did not fit into the existing item structures of the CPI were not introduced early.

In its report, the 1995 Boskin Commission produced its own estimate of the probable size of the error in the price inflation reported by the CPI. According to its computations, the CPI had been overstating the change in the cost of living by about 1.1 percentage points per year, this well before the 1990's (Boskin et Al., 1998). They reported that over a dozen years, the cumulative additional national debt from over indexing the budget by using the CPI could amount to more than 1$ trillion (Boskin et Al., 1998). According to them, the over indexing of government outlays and tax brackets had had a direct impact on the Federal Deficit and debt. Also, as suggested by Boskin and Jorgenson (1997) "because the CPI component price indexes are inputs into the national income accounts, an overstated CPI implies that real GDP growth has been understated".

The 1995 Boskin Commission not only quantified the extent to which the CPI had overstated inflation but quoting previous studies as well as their own, they specifically addressed

four "upward biases" present in the CPI methodology used by the BLS. These biases were called "upward biases" because they contributed in having the CPI "overestimate" the rate of inflation. Table 2 (at the end of the paper) lists all four "upward biases' and their respective estimates (as reported in Boskin et Al., 1998).

The first "upward bias" discussed and estimated corresponds to the "Upper Level Product Substitution Bias". The upper level product substitution bias occurs any time a product "among" a category is substituted for another. For example, this bias occurs when consumers purchase beef instead of chicken because a promotional sale results in a lower price for beef. This "substitution" by consumers, historically, had not been accounted for by the CPI. The 1995 Boskin Commission estimated this bias to be 0.15 of a percentage point. The second "upward bias" discussed and estimated by the Commission corresponds to the "Lower Level Product Substitution Bias". A "lower level product substitution bias" occurs any time a product "within" a category is substituted for another. For example, this occurs when consumers purchase Mackintosh apples instead of Red Delicious apples because a promotional sale results in a lower price for Mackintosh apples. This "substitution" by consumers was equally not accounted for by the CPI. The 1995 Boskin Commission estimated this bias to be 0.25 of a percentage point. Hence, both unaccounted "product substitutions" by the BLS in computing the CPI was estimated by the 1995 Boskin Commission to add-up to 0.40 of a percentage point (Boskin et Al., 1998).

The third "upward bias" discussed by the 1995 Boskin Commission was that of "outlet substitution". As reported by the 1995 Boskin Commission, in view of changes occurring in retail store formats, the BLS had been computing the CPI as follows: "Outlets are chosen and rotated every five years from a point-of-purchase survey, asking consumers where they purchase goods and services, with probabilities of outlet selection proportional to expenditures. There is thus approximately a 20 percent refreshing per year. The prices are collected and compared within outlets. No account is explicitly taken of substitution across outlet types, as might be expected with the evolution of retailing" (Boskin et Al., 1998). The 1995 Boskin Commission estimated that the outlet substitution bias corresponded to 0.10 of a percentage point. Interestingly, this bias is still an issues as illustrated by the title of a journal article by Hausman and Leibtag entitled "CPI Bias from Supercenters: Does the CPI know that Wal-Mart Exists?" which was published in 2009, hence, 14 years after the 1995 Boskin Commission (Boskin et Al., 1998).

The fourth "upward bias" discussed as well as estimated by the Boskin Commission was that of New Product /Quality Changes. This bias occurs when either new products and/or quality improvements impact prices. Here it is important to note that prices may not necessarily be lowered. Still, over time, prices generally will fall. This bias was estimated to be 0.60 of a percentage point and is larger than the upper level and lower level substitution bias combined.

As stated by Johnson, Reed and Stewart (2006) the "BLS has maintained that the evidence on quality bias and its direction are much less clear than for substitution bias". In view of new goods and how they would get accounted for by the BLS, they pointed out that new goods could enter the CPI computations in one of the following three ways: First, during repricing, if a sampled item was no longer available in the sampled outlet, the data collector would then "substitutes" to the most comparable item still remaining in that outlet and begin pricing it; Second, new goods could also enter the CPI sample through sample rotation (for which, as stated earlier, there are many critics that hold that the BLS has been much to slow in doing so); Finally, there was the case

of new goods that would not fit neatly into the existing CPI structure. These new goods would be introduced into the CPI only when a major revision of the item structure would occur. The 1995 Boskin Commission recommended that all four "upward biases" needed to be quickly addressed and remedied by the BLS. As will be seen later in the paper, many of the 1995 Boskin Commission recommendations would eventually be adopted by the BLS, especially the COLI framework which became the CPI's measurement objective in 1997. But as a reminder, our earlier identified "Adaptation Lag factor" was still occurring within the CPI .

THE 1999 CNSTAT PANEL

The fourth and latest professional review of the CPI, according to Reinsdorf and Triplett (2004), was done under the auspices of the 1999 Committee on National Statistics of the National Academy of Sciences (CNSTAT hereafter). Interestingly, while 34 years had elapsed between the 1961 Stigler Committee and the 1995 Boskin Commission, only 4 years separated the 1999 CNSTAT panel and the latter Commission. According to Reinsdorf and Triplett (2004): "The CNSTAT review of the CPI is most memorable for it partial retreat from the Stigler Committee's recommendation of the use of a COLI as the measurement concept for the CPI". The 1995 Boskin Commission had also recommended that the BLS adopt a COLI as the CPI's objective. This ambivalence by the 1999 CNSTAT panel is also observable in view of when the BLS adopted, in 1999, the use of geometric means that replaced the Laspeyres formula used in computing the CPI which required seasonality adjustments. This move to geometric means by the BLS accounted as well as helped correct substitution biases for most basic component indexes in the CPI. According to Reinsdorf and Triplett (2004) "Nevertheless, it was unclear to the panel that the geometric mean index would always be superior to the seasoned Laspeyres index". This ambivalence is again observable in the CNSTAT panel's views on the issue of outlet substitution bias: "The panel's review of the available evidence suggested that outlet substitution bias was significant enough to be a matter of concern, but they doubted whether researchers would be able to produce sensible, reproducible estimates for adjusting for quality differences between outlets. They therefore concluded that BLS had little choice but to continue research on the effects of outlet characteristics on prices" (Reinsdorf and Triplett, 2004).

According to Berndt (2006) the 1999 CNSTAT panel differed from the 1995 Boskin Commission in four important respects: first, it was much larger in that it consisted of thirteen members versus the 1995 Boskin Commission's five members; second, its composition was more diverse, including not only economists, such as in the case of the 1995 Boskin Commission, but equally a sociologist, a psychologist, and two statisticians; third, while the majority of the members of the 1995 Boskin Commission agreed with the appropriateness of a COLI framework for evaluating the CPI, by contrast, the 1999 CNSTAT panel "took nothing for granted, and started from scratch, vigorously arguing at considerable lengths among themselves on these and many issues". Finally, the CNSTAT panel took place during the late 1990s booming and exuberant economy which enabled it to operate at a more leisurely and academic pace than the 1995 Boskin Commission.

THE CPI'S MAJOR CONTROVERSY: THE ADAPTATION LAG FACTOR

It is our contention that the major controversy that has consistently "afflicted" the CPI corresponds to the unrelenting slow pace at which the BLS's has been updating the CPI in accurately reflecting changing consumer market conditions in the U.S. economy. This controversy, as may be observed from the preceding four milestone reviews, has been occurring since the mid-1940 or, for over 70 years. The BLS's "slow paced" adaptation to market changes, or what we term the "adaptation lag factor", has invariably as well as directly affected the CPI's accuracy in reporting price inflation of goods and services surveyed by the BLS. As an example of this "adaptation lag factor", the 1995 Boskin Commission reported that the time it took the BLS to include new products such as VCR's, microwave ovens, and personal computers in its sample of products for which it surveys consumer purchases and collects price data at retail outlets, was 10 years after they had penetrated the U.S. market. Most importantly, by then, their respective prices had fallen by 80% or more (Boskin et Al., 1998). This type of slow paced adaptation has been, in our opinion, a recurrent phenomenon with the CPI as computed by the BLS.

In our view, historically, two distinct groups, formed mostly of economists, have either supported or criticized the BLS's slow paced adaptation and actions in making the CPI reflect changing economic market conditions and consumer behavior responses: the first group has consisted of BLS "supporters" such as, for example, members of the 1942 Mitchell committee, Norwood (1964), as well as members of the 1999 CNSTAT panel, who, throughout the years, have mostly agreed and defended the BLS's slow pace in making the CPI better account for market and consumer behavior changes (hence, accurately measure price inflation). The Second group, we believe, has consisted of CPI "critics" such as, for example, the authors of the 1944 Meany-Thomas report, a majority of the members of the 1961 Stigler committee, MacDonald (1995), members of the 1995 Boskin CPI commission, and as shall be seen later in the paper, Leibtag (2006), Volpe and Lavoie (2007), Hausman and Leibtag (2009), and Greenlees and McClelland (2011), who all hold that the BLS is much too slow in updating the CPI's computational methodology, with the very important consequence that, as may be observed, the CPI has, over many decades, either been greatly "understating" or "overstating" price inflation of the goods and services it surveys.

THE CPI AND RECENT DEVELOPMENTS IN THE LAST TEN YEARS

As reported to Johnson, Reed, and Stewart (2006) subsequent to the 1995 Boskin Commission recommendations made public in their December 1996 report, the BLS made major advancements in addressing many of the Commissions' recommendations. Essentially, the 1995 Boskin Commission recommended that the BLS address the following four biases: Substitution (upper and lower level); new goods bias; quality bias; and outlet substitution bias. As mentioned previously, in 1997, the BLS first re-affirmed a "cost-of-living" index as an objective for the CPI's conceptual framework. In 1998, in view of the outlet substitution bias issue, the BLS changed its outlet rotation procedure to better account for the changes occurring in the retail sector of the U.S. economy. Hence, the CPI went from rotating 20 percent of the outlet sample each year to 25 percent so that the entire sample was rotated every 4 years instead of 5 years (Johnson, Reed, and Stewart, 2006). It addressed the substitution bias by way of accounting for price elasticity instead

of assuming it away. Specifically, as stated earlier, in 1999, geometric means were introduced in the CPI computation methodology which helped better account as well as reduce the upward bias resulting from upper level substitution. In view of the lower level substitution bias, in 2002, the BLS started producing a "Chained Consumer Price Index for All Urban Consumers" (C-CPI-U) which used a Tornqvist formula which replaced the Laspeyres formula which assumes zero substitution by consumers. The Laspeyres formula remained in use mostly for housing and medical care product categories. Excluding rent and owners' equivalent rent, Johnson, Reed, and Stewart (2006) report that "only one-seventh of the weights in the CPI still use a Laspeyres formula to calculate basic indexes". In view of the "quality bias", due to major advances in the field of hedonics, the BLS expanded the use of hedonic models to better account for changes in the quality of products (Berndt, 2006). In view of the "new goods bias", the BLS instituted procedures to introduce new goods more quickly into the CPI index by having more frequent updates to the item samples. Finally, in 2002, the BLS began updating expenditure weights based on consumer expenditure surveys every 2 years as opposed to roughly every 10 years in the past.

It may then be asked what impact did the revisions/changes made by the BLS to the CPI methodology have on its measure of price inflation. According to Berndt (2006) a positive impact had occurred: in June 1999, the U.S. General Accounting Office (GAO) initiated a study identifying methodological changes the BLS made to the CPI since the 1995 Boskin Commission. It asked the opinion of the four remaining members of the 1995 Boskin Commission as to how much of the bias in the CPI remained after changes were implemented by the BLS. The four former members of the 1995 Boskin Commission estimated that the changes brought by the BLS had reduced the annual upward CPI bias from 1.1 percentage points to between 0.73 and 0.90. It is, hence, puzzling to us to read what financial advisor Howard Simons wrote in 2004, a few years after the BLS had made major changes to its CPI computation methodology: "The CPI is a Laspeyres Index …. and is known to be an imperfect inflation measure. It ignores economic realities such as price elasticity of demand, substitution and technological improvement, so-called hedonic adjustments aside" (Simons, 2004). This quote would seem to suggest that some financial advisors were not fully aware of the BLS's latest efforts to improve the CPI's accuracy in measuring inflation. Interestingly, as reported by Berndt (2006), although the CPI is still acknowledged by some to be upward biased, hence overstate inflation, it has been found by others that in the case of some its goods and service categories, the CPI is likely to have been downward biased, hence, understating inflation. Berndt (2006) cites the following studies as examples: first, in a study by Nordhaus (1997) on the price of light, corresponding to the CPI's Energy product category, Nordhaus argued that, using CPI methods, its price increase would have been overstated by around 1.4 percentage points per year, this since about 1800; second, in view of Nordhaus' (1997) reported downward bias, Hulten (1997) argued that if this bias were true for the overall CPI as well as constant over time, then the implied standard-of-living for U.S. households in 1800 would have been implausibly low; third, and still in relation to Nordhaus'(1997) reported downward bias, a study by Gordon (2004) calculated that had the bias in the overall CPI been 1.4 percentage points annually since 1800, then the 1800 median household would have able to purchase only 1.3 pounds of potatoes per day, with nothing left over to pay for shelter, clothing or other goods. In addition, Gordon (2004) presented persuasive evidence that in the case of apparel, due primarily to the inability to link style changes reliably, there has been a downward bias over

time (and Berndt believes this downward bias to still be the case in 2006); finally, a study by Gordon and vanGoethem (2005) also documented a CPI downward bias for shelter, caused in part by the non response of tenants to CPI surveys who had moved just as rents were being increased. Interestingly, Berndt (2006), states that because of improvements brought by the BLS in its surveys since the mid-1980s, the CPI shelter downward bias is likely to be negligible. According to Berndt (2006) the BLS's major unfinished business was that of improving the accuracy of the CPI in accounting for the product category corresponding to medical care.

THE FOOD AT HOME CPI IN THE LAST TEN YEARS: THE LINGERING OUTLET SUBSTITUTION BIAS

Still, for many economists, the earlier reported important changes adopted by the BLS, beginning in 1997, so as to reduce the CPI's known upward bias in reporting price changes and inflation, were not enough nor done fast enough. Here we specifically review the "food at home" CPI literature since the publication of the CNSTAT panel recommendations in 2002. Leibtag (2006) reported that the CPI measure of food inflation was based on a selection of stores that had not been updated quickly enough to reflect the amount of food sold through big box stores. Leibtag (2006) estimated that annual food price changes as measured by the CPI had averaged an increase of 3% since the mid-1980s. Nontraditional food retailers such as Wal-Mart, Costco, and Target had gained more of the consumer for dollars since the mid-1990s: the share of sales going to traditional retailers such as conventional supermarkets had fallen from 82 percent to 69 percent by 2003. Previous studies had demonstrated that food prices at non-traditional retailers such as supercenters (Wal-Mart, Target) and wholesale clubs (Sam's, Costco) were on average 8-27 percent lower than at large supermarkets. However, these comparisons over store formats had not accounted for quality or package size. In conducting a study of dairy products and eggs, Leibtag (2006) found that dairy prices for similar package sizes were between 5 to 25 percent lower at nontraditional retailers than at traditional supermarkets.

Volpe and Lavoie (2007) investigated the competitive price effect of Wal-Mart supercenters on national brand and private label grocery prices in New England. As compared with conventional supermarkets, they found that Wal-Mart priced national brand goods 6 to 7 percent lower and that in the case of private label goods, prices were 3 to 8 percent lower. As compared to private label goods, supercenters had a greater impact in lowering prices on national brands. As national brands appeal to higher income consumers and that private label goods appeal to lower income consumers, they posited that what would explain their finding that Wal-Mart had lower prices for national brands than for private label goods was that private label goods, while universally cheaper than national brands, have higher markups. Conventional supermarkets, in facing Wal-Mart's lower price strategy, had increased their use of private label goods. Finally, Wal-Mart's lower prices were found to be most significant for both national brand and private label Groceries and dairy products and, least significant for national brand meat.

In an article entitled "CPI Bias from Supercenters: Does the BLS know that Wal-Mart Exists?" Hausman and Leibtag (2009) reported that the retail sector of the U.S. economy had witnessed immense changes that had begun before the new millennium: for example, they reported that Wal-Mart supercenters began selling food in 1998 and had became the largest U.S. grocery

chain by year 2002. Citing a study by Little (2004), they reported that supercenters (Wal-Mart, Kmart, Meijer, etc.), warehouse clubs (Sam's Club, Costco, and BJ's), and mass merchants (Wal-Mart, Kmart, and Target, etc.) which Little classified as "high-spend" expenditure stores (in contrast to "low and medium spend stores such as convenience stores) accounted, in 2003, for 24.8 percent of food expenditures in the U.S., with supercenters alone accounting for 45.6 percent of the category. In particular, Wal-Mart had become the largest supermarket chain in the United States, accounting for 14 percent of food sales while not being present, at the time, in many regional markets. When Sam's Club (owned by Wal-Mart) food sales were included, Wal-Mart's market share moved up to 18 percent. By year 2003, according to Little (2004), Wal-Mart's food sales, excluding Sam's Club food sales, had supermarket-related revenues approximately 51 percent larger than runner-up Kroger, and larger than the combined revenues of Albertsons and Safeway, respectively the third and fourth largest supermarket chains.

Hausman and Leibtag (2009) in analyzing the BLS procedure in computing the CPI, specifically took issue with what they believed caused two upward biases: First, they argued that the BLS's "linking procedure" used to incorporate new retail outlets in its sample of stores assumed that "quality adjusted" prices at Wal-Mart were exactly equal to prices at conventional supermarkets. Hence, the BLS procedure would, when including a Wal-Mart store in its outlet sample, link the lower Wal-Mart price to the higher conventional supermarket price in such a way as to remove any differences between both prices. According to them, the BLS in producing the CPI, made the implausible assumption that all price differences between supercenters and other stores were due to quality differences, nothing else. They argued that there was no empirical evidence that showed this to be the case. This bias is still in existence according to Greenlees and McClelland (2011). As they state: "The implicit assumption used in the CPI is that any cross-sectional differences in the prices charged in different outlets for the same item are attributable to outlet-related variation in "quality": stores offering lower prices may be less conveniently located, offer more limited product selection or hours of operation, and so on. Intuitively, in a state of static equilibrium in which outlets offer different prices there must be exactly offsetting differences in outlet quality. If not, one outlet would increase its share of the market". As pointed out by Hausman and Leibtag (2009), many past studies had shown that supercenters, mass-merchandisers, and warehouse clubs had in effect dramatically increased their market share and that these increases were not solely the result of "quality" differences.

Second, they argued that although the BLS updated its samples of store every four years as well as the goods and services in the market basket, they took issue with the "expenditure weights" which according to them were not updated quickly enough. It is important to note that although their article was published in 2009, it was based on a 2004 Conference presentation. Most interestingly, Berndt (2006) reported that the BLS had been, since 2002, updating expenditure weights based upon consumer expenditures every two years instead of the roughly every ten years in the past. Although Hausman and Leibtag (2009) acknowledged that the BLS had updated the outlet sampling procedure to a full rotation being done every four years instead of every five years, they simply considered the BLS "linking procedure" to be tremendously flawed.

Finally, Hausman and Leibtag (2009) in conducting a study to investigate the effect of supercenters, mass merchandisers, and wholesale club stores on food prices, found that Wal-Mart offered identical food items at an average price about 15 to 25 percent lower than traditional

supermarkets. Moreover, Wal-Mart's entry into a new geographic market created two different "price effects": first, a "direct" price effect by offering a lower price option to consumers; second, Wal-Mart created an "indirect" price effect by causing traditional supermarkets to lower their prices. According to them, the BLS "linking procedure" which cancelled out Wal-Mart's lower prices captured only the "indirect" effect, not the "direct" effect. Most importantly, Hausman and Leibtag (2009) concluded that annual food at home inflation was to high by 0.32 to 0.42 percentage points.

Greenlees and McClelland (2011), in studying the impact of the appearance and growth of new types of retail outlets on food prices found that the upward impact on prices from increased item quality had offset most, but not all, of the downward impact of lower priced outlets. Hence, in a strategy to protect their profit margins, they report that warehouse club stores, for example, had decided long ago to trade "low prices" for "increased sizes. They stated that their study's results offered by no means conclusive evidence of CPI bias but that "item quality" and "outlet Characteristics" were not negligible factors and warranted more research.

THE CPI AND MARKETING VALIDITY ISSUES

It may be noted that all of the CPI's four major biases (substitution, new products, quality changes, and outlet substitution) pertain to both the fields of economics and marketing. Substitution biases, whether "lower" or "upper" level, concern consumer behavior, an important field of study in marketing. The same may be said about "new products" and "quality changes". Outlet substitution specifically corresponds to the field of marketing channels with a special emphasis on retailing.

Validity in the social sciences such as for example Sociology, Psychology, Economics and Marketing corresponds to insuring that a measurement instrument measures accurately what it claims to be measuring, not something else (Zikmund and Babin, 2010). As mentioned previously, our inherited food at home survey instrument was composed of forty-one (41) food items. In September 2011, as our team began using this survey instrument, the following five (5) issues quickly came to our minds: these corresponded to the survey Instrument's: 1. Source; 2. Timeliness; 3. Number of items; 4. Item composition; and 5. Rural vs. Urban settings. In view of the first issue, we were never informed from what scientific basis/source the list of 41-food items originally came from. We assumed that it corresponded to a previous version of the list of food at home items surveyed by the BLS. We were told that, in 1977, the list had previously comprised 68 food-items, and that this number was scaled down to 48 items in 1983, this by another on-campus organization before it was transferred over to the now defunct federally funded agency. In view of the second issue, our inherited list of 41-food items dating back to 1983 was already 28 years old by 2011. More importantly, it was thought important to inquire about the present "Benchmark" in terms of the most scientific and currently available list of food items used by the U.S. Bureau of Labor Statistic in measuring, computing and reporting the Consumer Price Index (CPI) food at home price fluctuations. In view of the third issue, we found that, in 2012, the BLS verified the prices of 76 different food items in its report which, compared with our list of 41-food items, corresponded to 85% more food items surveyed. In view of the fourth issue, a quick perusal of the list of 41-food items revealed that 69% of the 41 items corresponded to "private-label

brands". We wished to verify if this proportion was similar or different to the list of food items used by the BLS in publishing the CPI. Finally, we desired to inquire if our list of 41-food items should not be "adapted" or modified when food price data is collected in a "Rural" area or setting such as our city of Plattsburgh. For example, the fixed Market Basket used by the CPI is designed to measure changes in the prices paid by "Urban" Households. To that effect, in 2012, the BLS, in computing the CPI each month, sampled 4,000 households located in 87 urban areas which kept records and reported prices they paid as well as collected price data from 26,000 retail establishments. Hence, adapting the Urban based food Market basket to produce a "Rural Food Market Basket" would seem to be a valid idea.

FUTURE RESEARCH

After gaining a solid understanding of the CPI's strengths and weaknesses, we foresee the following three research opportunities: First, we would want to assess our inherited 41-item food at home survey instrument's convergent validity this in view of the latest CPI food at home survey instrument published by the BLS. As mentioned earlier, since our inherited 41-item survey instrument dates back to 1983, this study would help us observe as well as compare the magnitude and direction of any resulting differences between both survey instruments. Using the latest (2013) CPI food at home list of items, we would make sure to collect supermarket/grocery data for at least three consecutive months. The second and third research opportunities would seek to replicate two parts of Volpe and Lavoie's (2007) research findings: First, we would want to assess if their findings of Wal-Mart's supercenters competitive price effects on grocery prices in New England would be similar in magnitude and direction in our upstate New York rural city of Plattsburgh (which also has a Wal-Mart supercenter); Second, we could conduct a study which would focus on the competitive effects of Wal-Mart supercenters on national and private-label grocery prices and establish if our findings based on Plattsburgh's Wal-Mart supercenter replicate those reported by Volpe and Lavoie (2007) in New England.

CONCLUSION

Although an important number of major recommendations made by the 1995 Boskin Commission have been adopted by the BLS in rendering the CPI more accurate in accounting for consumer behavior (both upper and lower level substitution, new product adoption, and quality changes) and Market changes (outlet Substitution), no other major revision by an appointed Commission or Committee has occurred since the 1999 CNSTAT panel. According to many economists, especially in view the food category and outlet substitution effects, the BLS still suffers from what we have termed the "adaptation lag factor". Studies undertaken within the last 10 years since the publication of the CNSTAT panel report such as those by Hausman (2003), Leibtag (2006), Volpe and Lavoie (2007), Hausman and Leibtag (2009), Greenlees and McClelland (2011) reveal that the CPI as well as the CPI-U food at home, still keeps overestimating price inflation by virtue of not correctly accounting the real impact of major changes that have occurred in the Retail sector of the U.S. economy, especially the major impact of Supercenters (such as Wal-mart) on retail prices as well as the still not properly accounted for

impact of "quality changes on prices**.** In view of outlet substitution, many economists believe that validity issues still plague the CPI as a measurement of price changes and inflation. Interestingly, outlet substitution bias, as estimated by the 1995 Boskin Commission corresponded to 0.10 percentage points out of the total 1.10 percentage points when all four sources of upward biases (upper level substitution, lower level substitution, new product/quality changes, and outlet substitution) are added. It actually was the lowest contributor to the CPI's upward bias, lower than upper level substitution which was estimated to be 0.15 percentage points.

In our opinion, what could best explain the CPI's long lasting controversy which we identified as an "adaptation lag factor" is what lies at the heart of the following quote by Simons (2004) who, almost 10 years ago, stated the following: "The CPI-U is subject to huge political pressures; government contracts, labor union agreements and escalators for social security are linked to its value". This idea of a political arena surrounding the CPI was equally expressed by Berndt (2006) who stated: "It is clear that the rise and fall of public interest in price measurement issues, including the Boskin Committee report and its legacy, needs to be interpreted in the political economy context of Congress and the White House attempting to deal with growing budget deficits…..". Reinsdorf and Triplett (2004) offered the following quote by Ostrander (1944) which stems from the time of the first professional CPI review by the 1944 Mitchell Committee and for which we believe best illustrates the impact of politics on the CPI: "It is not often that a price index, a tool of statisticians, becomes an object of political debate". Recently, a New York Post headline posted on the internet read "Ex-stats Insider: Time to Trash Outdated CPI". The New York Post Internet article by John Crudele stated that Keith Hall, former head of the Bureau of Labor Statistics, was of the opinion that the Consumer Price Index was broken and needed to be fixed (Crudele, 2013). We would not agree with "trashing" the CPI. Although the last ten years of food at home CPI research is characterized by mostly outlet substitution issues, as we have reported, there is sufficient controversy that maintains that not all goods and services surveyed by the BLS have suffered of an upward bias (for example, shelter and apparel). It is our opinion that too much effort has been invested over too many years by too many experts, this, in improving the CPI's accuracy, so as to abruptly put it to rest.

REFERENCES

Berndt, E. R. (2006). The Boskin Commission report after a decade: After-life or requiem? *International Productivity Monitor, 12*, 61-73.

Crudele, J. (2013, July 23). Ex-stats insider: Time to trash outdated CPI. *New York Post.* Retrieved from http://nypost.com

Boskin, M. J., Dulberger, E., Gordon, R., Grilliches Z., & Jorgenson D.W. (1998). Consumer prices, the consumer price index, and the cost of living, *Journal of Economic Perspectives, 12*(1), 3-26.

Boskin, M. J., & Jorgensen, D. W. (1997). Implications of overstating inflation for indexing government programs and understanding economic progress, Papers and Proceedings, *American Economic Review 87*(2), 89-93.

Franklin, A. W. (2001). The impact of Wal-mart supercenters on supermarket concentration in U.S. metropolitan areas. *Agribusiness, 17*, 105-114.

Goldberg, J. P., & Moye, W. T. (1985). The first hundred years of the Bureau of Labor Statistics. Washington, DC: GPO.

Gordon, R. J. (2004). Apparel Prices 1914-93 and the Hulten/Bruegel Paradox. Presented at the *CRIW conference on Price index Concepts and Measurements*, Vancouver, BC.

Gordon, R. J., & vanGoethem, T. (2005). A century of housing shelter prices: Is there a downward bias in the CPI?.

Greenlees, J. S. & McClelland, R. (2001). New evidence on outlet substitution effects in consumer price index data. *The Review of Economics and Statistics 93*(2), 632-646.

Hausman, J. & Leibtag, E. (2009) CPI Bias from supercenters: Does the BLS know that Wal-Mart exists?. In W. E. Diewert, J. S. Greenlees, & C. R. Hulten (Eds.), *Price Index Concepts and Measurement* (pp. 203-231). Chicago, IL: University of Chicago Press.

Hulten, C., R. (1997). Comment on Nordhaus. In T. F. Bresnahan & R. J. Gordon (Eds.), *The Economics of New Goods* (pp. 66-70). Chicago, IL: University of Chicago Press for the National Bureau of Economic Research.

Johnson, D. S., Reed, S. B. & Stewart, K. J. (2006) Price measurement in the United States: A decade after the Boskin report. *Monthly Labor Review* 10-19.

Leibtag, E. (2006). The impact of big-box stores on retail food prices and the consumer price index (CPI)', Economic Research Service by the USDA - Economic Research Report No. (ERR-33)

Little, P. (2004, June) Channel blurring redefines the grocery market. *Competitive Edge*, 34.

MacDonald, J. (1995). Consumer price index overstate food price inflation. *FoodReview*, *18*(3), 28-32.

Nordhaus, W. D. (1997) Do real-output and real-wage measures capture reality? The history of lighting suggests not. In T. F. Bresnahan & Robert J. Gordon (Eds.), *The Economics of New Goods* (pp. 29-66). Chicago, IL: University of Chicago Press for the National Bureau of Economic Research.

Norwood, J. L. (1964). The CPI controversy. *Labor Law Journal*, *31*(3), 131-138.

Ostrander, F. T. (1944). The Mitchell committee's report on the cost-of-living index: Comments. *The American Economic Review, 34*(4), 849 – 846.

Reinsdorf, M. (1993). The effect of outlet price differentials on the U.S. consumer price index. In M. Foss, M. Manser, & A. Young (Eds.), *Price Measurements and their Uses.* (pp. 227-254) Chicago, IL: University of Chicago Press.

Reinsdorf, M. & Triplett, J. E. (2004, June). A review of reviews: Ninety years of professional thinking about the consumer price index. Presented at the *CRIW on Price Index Concepts and Measurement*, Vancouver, Canada.

Schultze, C. L. & Mackie, C. (Eds.) (2002). At what price? Conceptualizing and measuring cost-of-living and price indexes. Panel on conceptual, measurement, and other statistical issues in developing cost-of-living indexes. Washington, DC: National Academy Press.

Simons, H. L.(2004, April). Can the CPI catch your eye?' *Futuresmag.com*, 50-53.

Volpe, R. J., & Lavoie, N. (2007) The effects of Wal-Mart supercenters on grocery prices in New England. *Review of Economics 30*(1), 4.

Wahl, R. W. (1982). Is the consumer price index a fair measure of inflation? *Journal of Policy Analysis and Management 1*(4), 496 – 511.

Zikmund, W. G. & Babin, B. J. (2010) *Exploring Marketing Research: Tenth Edition*. South-Western.

TABLES

Table 1 Consumer Price Index Expenditure Categories, July 2012 (Adapted from BLS Website)

Expenditure category	Relative importance Jun. 2012
	100.000
All items…….	14.208
Food…….	8.552
Food at home……	1.228
Cereals and bakery products……..	1.941
Meats, poultry, fish, and eggs……	0.889
Dairy and related products…….	1.266
Fruits and vegetables……..	0.941
Nonalcoholic beverages and beverage materials……...	2.287
Other food at home……	5.656
Food away from home…….	
	10.005
Energy……	6.108
Energy commodities……	0.214
Fuel oil……	5.794
Motor oil…….	5.612
Gasoline (all types)……..	3.897
Energy services……	3.045
Electricity…….	0.852
Utility (piped) Gas service ……...	
	75.787
All items less food and energy……...	19.763
Commodities less food and energy commodities…….	3.554
Apparel…….	3.173
New vehicles……	1.973
Used cars and trucks……	1.719
Medical care commodities……...	0.949
Alcoholic beverages…….	0.793
Tobacco and smoking products……	
	56.024
Services less energy services…….	31.411
Shelter……...	6.432
Rent of primary residence……	23.766
Owners' equivalent rent of residence……...	5.391
Medical care services……	1.605
Physicians' services…….	1.533
Hospital services……..	
	5.772
Transportation services…….	1.145
Motor vehicle maintenance and repair……...	2.407
Motor vehicle insurance…….	0.792
Airline fare……..	

Table 2: Estimates of Biases in the CPI Based Measures of the Cost-of-Living (Percentage Points per Annum) by the 1995 Boskin Commission (Source: Boskin et Al., 1998)

Source of Bias	Estimate
Upper Level Substitution Bias	0.15
Lower Level Substitution Bias	0.25
Subtotal	0.40
Outlet Substitution Bias	0.10
New Product/Quality Change Bias	0.60
Subtotal Total	0.70
Grand Total	1.10
Plausible Range	(0.80 to 1.60

ESTIMATING THE ECONOMIC BENEFITS A BUSINESS IMPROVEMENT DISTRICT WOULD PROVIDE FOR A DOWNTOWN CENTRAL BUSINESS DISTRICT

Inhyuck 'Steve' Ha, Western Carolina University
Sandra Grunwell, Western Carolina University

ABSTRACT

Business Improvement Districts (BIDs), as public-private partnerships have proliferated in the U.S. and globally as a means of revitalizing downtown central business. This study examined the potential of a BID enhancing a downtown central business district, in western North Carolina. The article reports on the process town officials activated to determine if a BID was right for their town and the outcomes of the process. The process included gathering information about downtown businesses via a survey to provide feedback to the town officials in charge of the proposed BID to maximize input for the plan and estimating the economic benefits the proposed BID would provide to both residential and commercial property owners and tenants. The business owner survey results aligned very closely with programs most BIDs provide to local businesses and property owners. Findings from the economic analysis revealed that a BID would be of value in enhancing retail sales growth and increasing property values. Based on the results of the study, town officials voted to approve the BID.

INTRODUCTION

Business Improvement Districts (BIDs), as public-private partnerships have proliferated in the U.S. and globally as a means of revitalizing downtown central business districts (CBDs) (Billings & Leland, 2009; Ewoh & Zimmermann, 2010; Grossman, 2008; Hoyt, 2006; Mitchell, 2001, Ruffin, 2010). "Described by The Economist as potentially 'the best hope of getting parts of America's cash-strapped cities working again,' business improvement districts (BIDs) are generating a great deal of excitement among city governments and urban policymakers around the world" (Ellen, et al., 2011, 1). The BID concept aims to promote and harmonize best practice in both urban management and tourism destination management. "Many of the characteristics that make a district a great place to visit – a variety of restaurants, cinemas, nightclubs, bars, cultural facilities, a walkable environment, attractive public places, and a feeling of safety – also make it a great place to live…as the visitor and residential markets tend to complement each other" (Ratcliffe & Flanagan,2004, 394).

According to the International Downtown Association (IDA) 2010 BID census report, all but two U.S. states have at least one BID, with an average of 20 BIDs per state. The median population size for U.S. cities with a BID is 102,804 people. BIDs exist across four continents and in 16 countries. The IDA report states that the total number of BIDs is 1,002 and that North Carolina has 51 BIDs, ranking 6th in the U.S. (Becker et al., 2011). The IDA currently functions as a central repository for information about BIDs for policy entrepreneurs around the globe (Hoyt,

2005). In performing the 2010 census and survey for the International Development Association (IDA), researchers needed to create a set of criteria for defining a BID. The criteria in the IDA survey and census included that the district be authorized by local and state government with a mandatory fee structure, be a public-private partnership where the government collects the tax but a non-profit management entity controls how much is collected and how the money is spent, and the district must perform traditional BID services such as cleaning, security and marketing (Becker et al., 2011).

PURPOSE OF THE STUDY

The study was initiated by the City of Asheville, North Carolina, the Asheville Downtown Association. To preserve this city and all the efforts that have gone into the period of revitalization since the 1990s, the Downtown Asheville Master Plan proposed a Business Improvement District (BID), similar to special districts that exist in many cities around the world. The desired goals for downtown Asheville focused on the continued healthy growth and evolution of the central business district, in order to continue increasing the strength of the regional economy. To reach the maximum potential of economic benefits, the BID should determine the exact needs of business and property owners and follow through on providing these requested services. Therefore, the purpose of this study was two-fold: (1) to gather information about downtown businesses to provide feedback to the town officials in charge of the proposed BID to maximize input for the plan and (2) to estimate the economic benefits the proposed Asheville BID would provide to both residential and commercial property owners and tenants.

LITERATURE REVIEW

There is not a universal definition for a business improvement district (Becker et al., 2011). This is because the very nature of a BID is to be flexible in order to serve each district as needed and desired by participants (Hoyt, 2005). "At their simplest BIDs are organizations entitled to levy an additional property tax within a specified area for providing a defined range of services or carrying out specific works" (Ashworth, 2003). BIDs are partnerships between business communities and local authorities to fund and develop projects that will deliver added value to the business environment (Grossman, 2012; Ewoh & Zimmerman, 2010). The property owners involved have a good deal of control over the amount of money collected and the services their money provides to them (Symes & Steele, 2003). In general, a BID consists of two or more business or property owners combining funds and creating programs designed to minimize obstacles to success and improve their profit, property value, and business or area improvement opportunities (Houstoun, 2004).

BIDs are usually governed by the city, but managed by a private, non-profit organization that is subject to a board made up of stakeholders, primarily business people and landowners, with some seats reserved for public officials, residents, community board members, and non-profit representatives. The municipal government typically collects the revenue and remits to the BID for services and projects supported by the property owners themselves (Briffault, 1999). There is no single approach to a BID and some may go far beyond the basic services. BIDs operate based

on the needs and desires of local conditions, community and stakeholders. Unlike taxes collected by the county, revenues collected from within the BID go directly back into the district for services and projects supported by the property owners themselves. The money collected from the BID property serves to provide supplemental services on top of that of local municipalities.

Having a reliable and predictable income year after year is an important success factor for business improvement districts. In most cases, revenue comes from charges on commercial property or directly to businesses only. Sometimes charges are levied on residential or other noncommercial properties as well. In North Carolina BIDs are called Municipal Service Districts (MSD). The frequency of North Carolina MSD tax rates, range between $0.0 and $0.6680 per $100 of assessed valuation for property within the district boundaries. The average tax rate is $0.1546 per $100. In North Carolina as of August 2010, the average rate was 0.1546, and the median charge was 0.14 per $100 valuation (North Carolina Department of Revenue, 2011).

"Property and business owners in urban contexts around the globe are using state authority to create a new form of government to protect their interests. With power to impose taxes and provide collective services, BIDs supplement publicly funded efforts to attract visitors and investors, enhance the pedestrian experience, and improve the city's ability to compete with regional office parks, shopping malls, and suburban living" (Hoyt, 2006, p. 221). BIDs provide services and improvements to boost business and property value; include a management entity to run the BID; and collect revenues through assessment-based annual mandatory tax on the properties within the BID boundary. BIDs focus primarily on creating clean, safe and attractive urban centers with downtown specific priorities that local governments often are remiss in providing due to financial restrictions. BID funding allows for the development and enhancement of streets, parks and buildings. There is little doubt that BIDs can give a strong competitive edge to towns and cities in the tourism market. As cited by Ratcliffe & Flanagan (2004), "a successful BID can increase an area's prosperity, attract inward investment and give a regional competitive advantage in terms of tourist destination management and visitor generation and spending."

In essence, a BID is a self-imposed way for downtown businesses and property owners to fund enhanced services or improvement projects within the district, using revenues generated by an assessment on real and personal property valuations in the district. Revenues created by the BID would be used for services and improvements over and above the level of municipal services already in existence. NC law requires that BID revenues be used only for services provided in the district, therefore, a BID would provide incremental services over and above the existing city services within the Asheville proposed district.

The NC MSD Act allows local governments, such as city council, to define a municipal service district for the purpose of levying an additional property tax amount on those properties within the MSD boundary. According to NC MSD Act, a city council may define any number of service districts in order to finance, provide, or maintain for the districts one or more of the following: beach erosion control, flood and hurricane protection works, downtown or urban area revitalization projects, transit projects, drainage projects, sewage projects, parking facilities, or watershed improvement projects (North Carolina General Assembly, G.S. 160A-536, 2011).

Across the world, a business improvement district is likely to devote itself to the advancement of the BID area's economy; provide services and improvements to boost business and property value; include a management entity to run the BID; and collect revenues through

assessment-based, annual mandatory tax on the properties within the BID boundary (Houstoun, Jr., 2005). BIDs seek to bring business into the downtown area instead of suburban shopping malls or retail centers (Symes & Steel, 2003). BIDs focus primarily on creating clean, safe and attractive urban centers with downtown-specific priorities that local governments often are remiss in providing due to financial restrictions. BID funding allows for the development and enhancement of streets, parks and buildings (Briffault, 1999). While street cleaning and maintenance as well as additional security are the principle goals of all BIDs, there is no single approach to a BID and some may go far beyond these basic services. BIDs operate based on the needs and desires of local conditions, community and stakeholders (Symes & Steele, 2003).

Unlike taxes collected by the county, revenues collected from within the BID go directly back into the district for services and projects supported by the property owners themselves (Briffault, 1999). The money collected from the BID property serves to provide supplemental services on top of that of local municipalities. Trash collection, for example, should continue as typically done by the local government. Services provided by the BID go above and beyond to produce greater results faster than those provided by local government through general taxation (Houstoun, Jr., 2005).

FACTORS FOR COMPARISION OF SIMILAR BIDs

When a city is considering a BID, studying other cities with BIDs with similar characteristics has value in being able to demonstrate what the BID did to support the area they serve. So for purposes of this study, five U.S. cities were selected with similar characteristics to Asheville that also had BIDs to determine the value of what these BIDs has accomplished for their town. Although a small city in terms of population, Asheville boasts a number of characteristics in common among larger cities: namely commitment to the arts and culture, historical preservation, a focus on livability for residents, the high number of retirees, active outdoor lifestyles and tourism draws, and the scenic mountain region. After selecting the cities with these similar characteristics, those with a BID service focus on supporting the arts and local culture, historical preservation and smart growth practices, clean, green and safety oriented programs, and business services were prioritized for comparison. The five cities selected for the final comparison case studies were Bozeman, Montana; Madison, Wisconsin; Portland, Oregon; Ann Arbor, Michigan; and Boulder, Colorado. Their BIDs were then also reviewed regarding their funding sources and their management structure, in addition to the services provided.

FUNDING

True to form, funding procedures are not the same for every business improvement district. However, in general, BID revenue comes from an assessment on property values within the district boundary, in addition to the property taxes paid to local government (Briffault, 1999).

In the International Downtown Association's 2010 census and survey (published in 2011), 95.9% of respondents stated that they obtained at least a portion of their revenue through property assessments. Other revenue sources were member dues (36.1%), contracts (41.2%), sponsorships (48.5%), development fees (21.6%), and funding from city general revenues (38.1%). Over half

(56.2%) responded that they had "other" methods of producing revenue, which include event revenue and vendor fees, alcohol sales at events, trust funds, sponsorships, charitable contributions, business license fees, grants, earned interest, participation fees, rental income, parking fees, transportation fees, and more. About half (55.9%) of respondents calculated their assessments based on value of real estate as calculated for tax purposes, 1.8% based assessments on sales tax, 12.2% on square footage, 4.5% on linear front footage, and 25.7% on "other" bases (Becker, et al., 2011).

In most circumstances, the charges levied to BID property owners are treated like taxes in that failure to pay results in legal action such as a fine, a lien against the property, or a delinquency sale. These legal implications indicate the role government plays in controlling BID financing. For this reason, BIDs almost always count on local government tax collection services to bill the property owner and collect the BID's revenue (Briffault, 1999). Although most revenue comes from these levied fees, BIDs are not limited to this sole source of income.

Some BIDs do receive other financial support in addition to the revenue from assessments on property value. Tax-exempt property owners including government, non-profit and religious organizations operating within the district may provide voluntary funding to the BID. BIDs are eligible for economic development grants from federal and state agencies. They are also able to collect interest income and proceeds from bonds backed by revenue from the district. Revenue may come from fees or charges for use of district facilities, or managing publicly owned facilities (Briffault, 1999).

In most cases, revenue comes from charges on commercial property or directly to businesses only. Sometimes charges are levied on residential or other noncommercial properties as well. The charges to these entities may be lower than the charges to commercial entities and property (Briffault, 1999). In the United States tend to be much lower than in other countries. Common charges here may be 15 percent of property tax, coming out to as little as 50 cents per day, or 10-15 cents per square foot (Houstoun, Jr., 2005). Other studies show assessments in the United States are often below 10 percent of property tax. A 1995 Pittsburgh Downtown Partnership study of twenty-three BIDs showed charges ran from six to eight cents per square foot and a more recent study showed charges to be 10 to 12 cents per square foot (Briffault, 1999). This also rings true in North Carolina where, as of August 2010, the most common rate was 0.10, the average rate was 0.146, and the median charge was 0.14 per $100 valuation (North Carolina Department of Revenue, 2011).

Having a reliable and predictable income year after year is an important success factor for business improvement districts. This is the advantage of the compulsory tax model. BIDs in the United States who raised their revenue primarily through voluntary contributions spent up to half of their management time fundraising instead of using that time to provide services and programs for long term success (Lloyd et al., 2003). Funding models where the BID relied most heavily, if not only, on voluntary funding models have been shown to fail. Typically, only a small number of businesses or individuals will contribute. With only a few carrying the cost for all, the non-contributors have no incentive to participate and those who do contribute eventually become fatigued of supporting the entire district (Houstoun, Jr., 2005).

Compulsory assessment provides long range, stable, and secure funding for the BID to maintain its services and programs (Briffault, 1999). This background research strongly indicates

that BIDs are most successful when operating with a for-profit, non-voluntary payment attitude and structure. This enables BID management to produce positive results, ultimately resulting in the approval of local property owners and tenants.

SERVICES

The services most commonly provided by BIDs throughout the world include capital improvements (such as street lighting and greenery, sidewalks and curbs, bus shelters, trash bins, wayfinding signage), consumer marketing (including events), economic development (incentives or loans to bring in and help expand business), maintenance (such as street and sidewalk cleaning, landscaping, graffiti removal), policy advocacy (including lobbying government for district commercial interests), security, social services (including job training, homeless services and youth activities) and transportation (including parking) (Hoyt, 2005, Mitchel, 1999). Most, if not all, BID services could likely be put into one of these categories.

By surveying business owners, residents, and visitors (both locals and tourists alike), BIDs are better able to provide the services and programs that will lead to the most positive outcome for stakeholders. For the purpose of this report, we use fewer and broader service categories, dividing the background research into physical improvements, business services, and supplemental municipal services. For the purpose of this report, fewer and broader service categories were used, dividing the background research into physical improvements, business services, and supplemental municipal services.

PHYSICAL IMPROVEMENTS

Physical improvements to the BID, also considered capital improvement projects, include long term projects and any major improvement project varying from street repair or paving, sidewalk or curb repair or creation, landscaping including new trees, flowers and plants, and street furniture such as benches, shelters, kiosks, lamps and hydrants (Briffault, 1999). In some states, BIDs are able to finance capital improvements by floating bonds, using their own income to help leverage state and local funds (Houstoun, Jr., 2005). These items are typically big projects that the city government may not be able to afford on their own but the district can help pay for and benefit from.

SUPPLEMENTAL SERVICES

Most BIDs seek to primarily boost and go beyond the municipal services provided by the government. This includes sanitation, security and maintenance (Briffault, 1999). Speaking to the fear that a BID may replace or reduce government provided services, it is important to point out that BIDs do not provide the same services as the municipality. Marketing is one example (Houstoun Jr., 2005). Additional sanitation or security services even serve a marketing and promotion purpose when BID workers dressed in uniforms to highlight and promote the efforts of the BID (Briffault, 1999).

SOCIAL SERVICES

While not very common among business improvement districts, social services are highlighted here due to the demographic profile of downtown Asheville and the Asheville community's involvement and commitment to providing such services. When a BID does provide social services, they are often a small portion of their programs and typically involve the homeless. These include shelter, food, employment and training opportunities, or referral services. In most cases, these programs come from a desire to maintain public order and the appearances of the district (Briffault, 1999).

BUSINESS SERVICES

The services provided and programs implemented by a business improvement district are all intended to boost the business within the district. BIDs may assist with finding renters for unoccupied space or buyers for buildings for sale, financing for a new business, recruiting new businesses or helping balance the business mix within the district. Some may provide grants or loans to help businesses improve the façade of their building. Those services specifically targeted to business include promoting and marketing products and services provided by businesses within the district, recruiting and retention of businesses in the district, and attracting visitors, consumers and tourists to the district. (Briffault, 1999).

In addition, recruiting businesses into the BID may be an essential part of the BID's services, depending on the priorities determined by the stakeholders in the area. A great way for BIDs to strengthen current business and recruit new business is by partnering with economic development agencies – including those focused on broader areas such as the city, county, region or state (Houstoun, Jr., 2004). This is another example of the public-private partnership so essential to business improvement districts.

SMART GROWTH

Numerous BIDs across the United States have adopted a Smart Growth plan, with mixed use development and pedestrian friendly goals in mind that promote recreation and culture, as well as unify the vision for the city. Boulder, Colorado was the first city in the nation to proactively advocate fundraising for the purchase of green space in areas surrounding the city, and began addressing traffic concerns many years ago (Benfield, et al. , 2001). Historic preservation is essential to maintaining the unique sense of place that is found within all vivacious downtown communities. Both, Boone, NC and Charlotte, NC BIDs have historic preservation committees, with emphasis placed in various areas.

MANAGEMENT STRUCTURE

The management structure for a business improvement district is an example of the public-private nature of BIDs. Business improvement districts are usually governed by the city but managed by a private, non-profit organization that is subject to an advisory board or board of

directors made up of stakeholders. The governance structure usually specifies specific formal roles for the city and the property owners. It is this advisory board and administrative body that is referred to as "The BID", since the board and administrative organization make recommendations and carry out the services and programs performed (Briffault, 1999). One study found that residents and government agencies contributed to the formation of a BID but that once established, programs and services efforts were typically led by commercial property and business owners (Hoyt, 2005).

Just like with initial formation of the BID, approval by municipal government (the ultimate governing body for the BID) is required for a change in boundary, assessment fees, or bonded debt for capital projects. However, in general, the BID management association and the board typically see little interference or control exerted from the municipality (Briffault, 1999). The municipal government typically collects the revenue and remits to the BID. Even though the association serves as a management entity and not the governing authority (this, again, is the city), it is still the management association that is usually considered responsible for policy and fiduciary day-to-day functions. The management association is then likely to contract out the services the BID chooses to undertake, such as administrative, security, sanitation and landscaping or maintenance services. This is especially true for BIDs with smaller budgets and smaller management associations. Often, the contracted service provider was a proponent of the BID before formation (Briffault, 1999).

Briffault (1999) also found that advisory and administrative boards were primarily made up of businesspeople and landowners, with some seats reserved for public officials, residents, community board members, and non-profit representatives. Boards are either appointed by the government or elected by the district stakeholders – but appointment is much more common

DATA AND METHODOLOGY

This study uses data from the U.S. Census and ESRI to create a portrait of the demographic background of the central business district in Asheville. Census data at the census tract level provides information on residents inside the central business district. It is important to note that the census tract is 35% smaller in area than the central business district. Therefore, ESRI was used to create a custom polygon of the central business district to show the differences between the census tract and the CBD. ESRI reports also provide projected demographic information for 2010 and 2015, based on the 2000 Census, including quarterly information on population and households from January 2009 to October 2010, a market profile, and expenditures information from 2000 and projected to 2015.

In order to gather information about downtown businesses and provide the City of Asheville, Asheville Downtown Association, and the Downtown Master Plan Commission with feedback from business owners and tenants, the Asheville Downtown Business survey was created and disseminated via email. A total of 100 surveys were completed.

To estimate the economic impact of a new BID in Asheville, an input-output model was constructed. The researchers utilized the IMPLAN (IMpact Analysis for PLANing, Minnesota IMPLAN Group, 2007) software input-output model and database to construct a basic input-output

model. The input-output model is useful for estimating the economic impact and understanding how the impacts ripple throughout an economy.

Direct dollars spent for goods and services identified within the IMPLAN model as items that are available from within the regional or local economy are traced by an input-output analysis as secondary impact dollar spending. Secondary impact dollars accumulate as a result of both indirect and induced effects. Indirect effects are secondary impacts that result from businesses that make expenditures in order to replenish goods and improve services that have been purchased by direct (initial) impact expenditures. Induced effects are secondary impacts resulting from an increase in household spending by employees who are hired, or current employees paid to work longer hours, to provide goods and services being purchased.

Estimates of secondary impacts are based on a multiplier effect, an economics principle widely used to calculate spending that takes place as a result of the "ripple effect." The concept is that every dollar received by business owners and employees is re-spent, multiplying the initial sales and generating revenues in other sectors of the local economy. IMPLAN estimates the magnitude of both primary and secondary impacts for each industry, which is so-called "multipliers."

It should be noted that a portion of direct and secondary dollar spending goes for goods and services that are not purchased in the local community as well as to pay taxes. Money used to purchase items that are not available in the local community and money used to pay state and federal taxes leaves the local economy, and so do not continue to circulate within the local economy.

RESULTS

DOWNTOWN BUSINESS SURVEY

Survey results indicated that businesses currently located in downtown Asheville are committed to operating in a downtown location, 67% had been in operation in their current downtown location for 5 to 20+ years, 30% had been in operation in their current location for over 15 years. Almost all respondents (93%) were satisfied with their current location, and the vast majority (90.9%) had no plans to relocate. Of those who did plan to relocate, all except one, planned to relocate in the downtown Asheville.

Most business owners rent their spaces, demonstrating that those operating businesses downtown are not the ones owning and pay taxes on the property. This causes some concern for both owners and renters when discussing an increase in taxes to fund a BID. Most likely, the owners would transfer the cost onto the renter as part of the terms of their lease, given that the renter will receive the most direct benefit from services provided on a daily basis by the BID – although the owner will ultimately benefit from the expected increase in property value.

Respondents were asked to respond to statements as to what were to most competitive traits downtown Asheville had to offer over other area destination shopping districts. The most competitive traits from the business respondents' perspective were character/sense of place, location, the trend to 'shop local' and the promotion of that trend, quality of products/services, and customer service. When respondents were asked to what degree they were experiencing any of

the challenges provided in a list in their downtown business, the biggest challenge was seen as insufficient parking, followed by street people or panhandlers, cost of rent, expense of employee wages and benefits, and street closure for events.

To determine the general attitude of local businesses the respondents were asked to respond to a list of statements about their perception of downtown Asheville. The results showed respondents felt downtown Asheville was an excellent place to have a business. The results also showed that local business respondents sought ways to be cooperative rather than competitive with one another, including directing consumers to other downtown businesses, seek ways to cooperate with complimentary downtown businesses, try to buy products /services downtown, and felt the existing business mix was helpful to their own business. These comments suggest that downtown business owners feel positive about operating a business in downtown Asheville, and feel positive about other downtown businesses.

However, the results also showed that respondents were largely unhappy with local municipal services as they are currently provided. Respondents disagreed with all statement involving maintenance, police protection and safety, and municipal service; such as sidewalk and street maintenance is outstanding, local services are worth the level of taxation, locale waste management service is outstanding, feeling safe downtown, and that local police protection is outstanding. Revealing respondents were largely unsatisfied with these local municipal services as they are currently provided. While the city may not like to hear these results, it does provide a snapshot of what the BID could offer in addition to local services in order to better please downtown businesses and in attraction visitors to the city.

Respondents were also asked to respond to statements designed to determine what they considered their most valued services and improvements. The most important services or improvements selected by the majority of the respondents were: sidewalk and street cleaning, clearing sidewalks when it snows, additional parking, marketing, business and economic development, and additional security or police. These responses align very closely with the programs most business improvement districts provide to local businesses and property owners. This should help support the effort to implement the Asheville BID, and it is important that these responses be taken into consideration when planning the BID.

Respondents were asked their total sales for the year 2010. Almost half (45%) of respondents had total sales between $250,000 and $1,000,000 during the 2010 year. When asked what percent their total sales increased or decreased in 2010 compared to 2009 total sales, 49.5% had an average rate increase of 16.1%; whereas 24.7% had an average rate decrease of 20.6%, and 25.8% stated their total sales stayed the same. When asked what percent they expected their total business sales to increase or decrease in 2011 compared to 2010, 63.5% expected an average rate increase of 14.4%, whereas 10.4% expected an average rate decrease of 11.3%, and 26.0% expected their total business sales to stay the same.

ECONOMIC BENEFITS ANALYSIS

It is expected that implementation of a BID in downtown Asheville would have a positive economic impact on property values and retail sales within the district lines.

It should be noted that there are very few empirical studies focusing on the impact of a business improvement district on property values and retail sales. Most studies and reports on BIDs reiterate this lack of findings and tend to focus on the attitudes of business and property owners and the completed projects as the BID's accomplishments. It is difficult to estimate the economic benefits of a BID before implementation, in part because the magnitude of these benefits depends on the success of the BID and its programs.

In the central business district, which is currently the area of downtown Asheville being considered for the BID, there were a total of 1,392 property tax parcels in 2010-11. Of these, 1,257 properties (90.3%) are non-exempt properties. The total appraised values of real properties from 2010-11 is $1,135,425,387. The total value of tax-exempt property values, however, is $468,865,900. This leaves $666,559,487 (58.7% of all property tax value) in value from non-tax exempt property. If the BID tax rate is $.10 per $100, then estimated tax revenue for the BID is $637,411.69. If personal properties are included, the estimated tax revenue for the BID is $795,438.05. If $.10 is levied and spent in the district, then it is estimated that 16.1 jobs are created and maintained annually.

Given available data and estimates, such as both commercial and residential property values using parcel data in the Asheville CBD and the surrounding 28801 zip code, census tract data, statistical analysis results based on the property values, and statistical estimates from the existing literature, we can expect at least two percentage points positive in property values annually for the next several years with the implementation of the BID.

The Asheville Business Survey indicates that local businesses expect an increase in total sales by 14.4% this year. Trends in retail sales, gathered from NC Department of Commerce, indicates Buncombe County experienced a 7.5% increase in total sales last year. The population in the Asheville's CBD grows about 1.0% annually estimated from the ESRI data, and visitor spending has increased by 1.6% annually for the previous six years according to the Asheville Area Tourism Research published by Buncombe County Tourism Development Authority.

With the implementation of the Asheville BID in the central business district, the estimated average annual growth in retail sales is 5.3% annually in addition to normal growth rates without a BID. As previously stated, property values will increase at least 2% annually on top of growth without a BID.

The expectation for positive impact is due to the uniform nature of municipal services provided by local government, when some areas may demand more than the local municipal services can supply. The positive economic impact, therefore, comes from the additional services the BID provides to meet the excess demand. To reach the maximum potential of economic benefits, the BID should determine the exact needs of business and property owners and follow through on providing these requested services.

CONCLUSIONS AND RECOMMENDATIONS

Based on the results of this study, it was recommended that Asheville form a business improvement district as proposed in the Asheville Downtown Master Plan. This recommendation was based on the extensive examination of business improvement districts, Asheville's characteristics, specific case studies similar to Asheville, a survey of Asheville central business

district property and business owners, and an economic benefits analysis on impacts, property values, and retail sales.

Secondary research indicates that business improvement districts are typically successful, even though the measures of that success are not often empirical. Primary research in the form of an economic benefits analysis shows that a BID would provide an additional flow of income that, if used effectively, could improve the impressions and perceptions of business and property owners in the central business district of downtown Asheville. This analysis shows that in addition to positive perceptions, the BID would also likely provide a boost to property values and retail sales although this might be difficult to prove after the implementation of the BID due to uncontrolled factors.

As previous BID studies have found, there are a variety of management and funding structures used internationally. Based on previous successes and failures of other BID structures, it is recommended that the Asheville BID consider a compulsory tax as their primary BID income source. Voluntary donations, grants and fundraising as a primary funding structure tend to take too much time away from the programs and services a BID needs to focus on in order to be the most successful.

The city typically collects the revenue and sets it aside for the purpose of the BID. Cities will often provide a liaison to the BID, often through their chamber of commerce or economic development office. Many BIDs then choose to contract with a management organization that will run the operations of the BID. They will often contract out for services such as garbage collection, cleaning, and beautification. BIDs may also choose to hire their own staff to manage operations and execute programs and services, or they may do a combination of direct hiring and contracting. In some cases, if a downtown association is already established, this is an easy transition into providing BID programs and services effectively, rather than starting from scratch. Although aware of concerns about the Asheville Downtown Association becoming the management organization for an Asheville BID, the researchers believe it would be a smooth and effective transition if ADA were to take on management of - or at least partnership with - the new BID. For example, if ADA did not take on the management role, the BID could partner with ADA to continue to provide the many annual events that promote downtown.

BIDs also commonly have a Board of Directors made up of stakeholders in the district. Typically there are seats available to represent both large and small property owners, business owners who rent their space, large non-profit organizations, and residents in the area. BID managers usually look for as much diversity as possible in their board. It is recommended that Asheville follow this trend should the BID be implemented. This may be an especially helpful component when trying to constructively engage and work with dissenters or those who are unsure how they can benefit from the BID.

As previously discussed, it is extremely important to the success of a BID that the services most valuable to the stakeholders are provided by the BID. Otherwise, participants will be unsatisfied and may repeal the BID. It is recommended that the Asheville BID, perhaps even prior to implementation, do extensive surveying of all businesses and residents as well as property owners, to determine what services they feel are highest priority and worth the cost. With those results, the BID can ensure the programs and services delivered to the district result in stakeholders believing they are receiving value from the small increase in taxation. It is also important for these

services to be implemented as soon as possible at the highest quality possible, in order to cement positive perceptions of the BID.

The services provided by the BID should also be supplemental to the services provided by the city. Often, a BID covers service needs above and beyond what the municipality can provide. This is where BID revenues and services can provide a positive impact. If the Asheville BID needs additional trash and recycling pick-up or street cleaning that goes beyond the needs of other service areas, the municipality is not likely to be able to provide this extra service. Therefore the BID can administer, or contract with an agency to administer, supplemental services that accommodate the needs of the Asheville BID area specifically. These services do not take the place of the municipal services already provided, and it is recommended that the BID take care to ensure the municipality does not slacken its efforts due to knowing the BID provides these supplemental services.

When the formation of a BID seems to effectively lessen problems such as crime within the district area, sometimes those problems are simply being shifted from within the BID boundary to outside of it. The benefit of the BID becomes the detriment of the surrounding area, which ultimately is not good for the BID either. Property values can sometimes go up in a BID but only in relation to the property values going down in the surrounding area. It is recommended the Asheville BID, therefore, hold onto awareness of this possibility and attempt to help eradicate these issues rather than simply push out crime and other problems to outside the boundary lines.

Assessing performance is varied among BIDs. Since it is difficult to demonstrate that positive changes within a BID are directly caused by the BID itself, most do not provide specific data assessments on a regular basis. Some BIDs do provide a list of accomplishments for the area, including increase in tourism, increase in sales, garbage collected, projects completed, and decrease in crime rates. This may be on an annual basis or may not occur at all. It is recommended that the Asheville BID, if implemented, perform a regular survey of business and property owners, residents and also visitors to the BID area. If possible, this survey should be done annually or bi-annually. It is also recommended that the Asheville BID consider keeping a close record of the monies raised and spent within the BID for accountability purposes with stakeholders.

The researchers believe time is of the essence in forming an Asheville BID. It is rarely easy to convince property owners to pay additional taxes. It may be even more challenging given the current economy. Yet the additional income to the central business district may be all the more important due to the current local, state, and national economic conditions. If the advocates for a BID can frame their promotion of a BID in a way that makes sense and is compelling to property owners they will be more likely to agree. It will be very important in this process to promote feasible plans for the improvements and services which are most important to property and business owners. On October 9, 2012 the Asheville City Council approved the Asheville BID that was initially proposed in the 2009 Asheville Downtown Master Plan.

REFERENCES

Ashworth, S. (2003). Business improvement districts: The impact of their introduction on retailers and leisure operators. *Journal of Retail & Leisure Property*, 3(2), 150-157.

Becker, C., Grossman, S., & Dos Santos, B. (2011). *Business improvement districts: census and national survey*. Washington, D.C., International Downtown Association.

Benfield, F., Terris, J., & Vorsanger, N. (2001). *Solving sprawl: models of smart growth in communities across America.* New York: National Resources Defense Council.

Billings, S., & Leland, S. (2009). Examining the logic behind the self-help, self-taxing movement: business improvement district formation. *Public Budgeting & Finance*, Winter 2009, 108-124.

Briffault, R. (1999). A government for our time? Business improvement districts and urban governance. *Columbia Law Review,* 99(2), 365-477.

Ellen, I., Schwatz, A., Voicu, I., Brooks, L., & Hoyt, L. (2007). The impact of business improvement districts on property values: evidence from New York City. *Brookings-Wharton Papers on Urban Affairs.* [On-line]. Available at http://www.jstor.org/ stable/25067439. [Retrieved September 6, 2012], 1-39.

ESRI 2011. ArcGIS Desktop: Release 10. Redlands, CA: Environmental Systems Research Institute.

Ewoh, A., & Zimmermann, U. (2010). Public-private collaborations: The case of Atlanta Metro Community Improvement District Alliance. *Public Performance & Management Review,* 33(3), 395-412.

Grossman, S. (2008). The case of business improvement districts: Special district public-private cooperation in community revitalization. *Public Performance & Management Review*, 32(2), 290-308.

Grossman, S. (2012). The management and measurement of public-private partnerships: Toward an integral and balanced approach. *Public Performance & Management Review*, 35(4), 595-616.

Houstoun, Jr., L. (2004). Business improvement districts. *Economic Development Journal*, 3(3) 49-54.

Houstoun, Jr., L. (2005). Business improvement districts reach Europe. *Economic Development Journal*, 4(3), 15-23.

Hoyt, L. (2005). *The business improvement district: an internationally diffused approach to revitalization*. Washington, D.C.: International Downtown Association. [On-line]. Available at http://web.mit.edu/dusp/dusp_extension_unsec/people/faculty/lhoyt/ Hoyt_IDA.pdf. [Retrieved September 6, 2012], 1-65.

Hoyt, L. (2006). Importing Ideas: The transnational transfer of urban revitalization policy. *International Journal of Public Administration*, 29, 221-243.

Lloyd, M., McCarthy, J., McGreal, S., & Berry, J. (2003). Business improvement districts, planning and urban regeneration. *International Planning Studies*, 8(4), 295-321.

Minnesota IMPLAN Group, Inc. (2007). IMPLAN professional software 2.0.

Mitchell, J. (1999). Business improvement districts and innovative service delivery. Grant Report. *The Pricewaterhouse Coopers Endowment for the Business of Government*. [On-line]. Available at http://jj0955.com/PdfFiles/MitchellBusinessImprovement Districts.pdf. [Retrieved September 6, 2012], 1-36.

Mitchell, J. (2001). Business improvement districts and the "new" revitalization of downtown. *Economic Development Quarterly*, 15(2), 115-123.

North Carolina Department of Revenue (2011). *Revenue's preliminary report on property tax rates and latest year of revaluation for North Carolina counties and municipalities: fiscal year 2009-2012, August 2009.* [On-line]. Available at http://www.dor.state.nc.us/ publications/2009-10taxrates_prelim.xls. [Retrieved July 7, 2011].

North Carolina General Assembly (2011). *G.S. 160A-536. The Municipal Service District Act*. [On-line]. Available at http://www.ncleg.net/EnactedLegislation/Statutes/HTML/ ByChapter/Chapter_ 160A.html. [Retrieved July 7, 2011].

Ratcliffe, J., & Flanagan, S. (2004). Enhancing the vitality and viability of town and city centres: the concept of the business improvement district in the context of tourism enterprise. *Property Management*, 22(5), 377-395.

Ruffin, F. (2010). Collaborative network management for urban revitalization. *Public Performance & Management Review,* 33(3), 459-487.

Symes, M., & Steel, M. (2003). Lessons from America: the role of business improvement districts as an agent of urban regeneration. *The Town Planning Review*, 74(3), 301-313.

U.S. Census (2000). Summary File 1 (SF-1) 100-Percent Data, U.S. Department of Commerce.

EFFECT OF GOVERNMENT DEFICIT SPENDING ON THE GDP IN THE UNITED STATES

Morsheda Hassan, Wiley College, Marshall, Texas
Raja Nassar, Louisiana Tech University, Ruston, LA
Chang Liu, SouthernWest University, China

ABSTRACT

The relationship between government deficit spending and the growth domestic product is of extreme importance for economic policy making, especially in times of economic downturns as has been experienced in the US and around the world in recent years. The literature is mixed on this issue. There are studies arguing that deficit spending has an adverse effect on the GDP by way of increasing the interest rate and hindering business investment. Other studies argue for deficit spending at a time of recession as being beneficial in that it spurs demand and has no effect on interest rate.

It is important to look at data in order to determine if deficit spending has an effect on GDP in the presence of control variables such as interest rate, unemployment, and inflation which may have an effect on the GDP.

In this study, we analyze data from the US and develop a time series model showing the relationship between deficit spending and GDP. Results revealed that government deficit spending had a negative effect on GDP. Inflation rate and interest rate had no effect on the GDP. Only unemployment had a negative effect on the GDP in the presence of deficit spending. It is interesting to note that GDP was cointegrated (having a long-run equilibrium relationship) with unemployment rate, interest rate, and inflation rate.

INTRODUCTION

Deficit spending by a government and how it relates to the economy as measured by the GDP is of fundamental importance in shaping economic policy for a country, especially at a time of economic downturn. It is of importance to determine if deficit spending would have an influence on economic growth. The literature is mixed on this issue. Some economists are of the opinion that deficit spending has a negative effect on the economy in that it increases interest rate which leads to a decrease in investment. Others argue that deficit spending has a positive effect by increasing demand. On the other hand, there are those that argue for no effect on the economy.

If deficit spending has a negative effect on growth, then fiscal austerity would be legitimate as a remedy for an economic downturn. On the other hand, austerity measures would be detrimental to the economy if deficit spending is the right stimulus for growth.

From the arguments above, it is clear that empirical studies to determine the relationship between government spending and economic growth are of utmost importance. In this study, we employ time series analysis techniques to investigate the relationship between federal deficit spending and economic growth in the United States over the time period 1930-2010.

RELEVANT LITERATURE

Kiani (2007) showed that there was a positive effect of budget deficit in the US and long term interest rate. Also, a positive effect of budget deficit on interest rate in the United States was reported by Feldstein (1986) and Holster (1986). Posner (1987), Krueger (2003), and Macao (2003) argued that an increase in interest rate would cause a decrease in investment, which would decrease economic growth. On the other hand, studies by Makin (1983), Plosser (1982), and Evan (1985) showed no evidence of a relationship between budget deficit and interest rate.

Cebula (2008) provided evidence through co-integration analysis and an error correction model, indicating that federal deficit in the US, over the period 1973-1996, and interest rates on high-grade tax-free municipal bonds were positively correlated. An increase in budget deficit is accompanied by an increase in real interest rate. Also, there was some indication that high interest rate has a positive effect on federal deficit. This indicates a bidirectional relationship between the two variables. It is difficult, however, to infer cause and effect from a correlation.

Collins (1999) presented coefficients of correlation (on data between 1944 and 1994 in the US) between deficits and stocks and bonds as well as data on deficits and investment and interest rates. Results were not consistent with the argument that deficits cause an increase in interest rate and a decrease in GDP growth, investment, and stock performance.

Nikannen (1978) reported that budget deficit led to an increase in government spending, but had no effect on inflation over the period 1947-1976 in the United States

Pollin (2012) concluded from his study that the US government deficit related to the 2009 economic stimulus did not cause an increase in interest rate or inflation.

Siklos (1988) using spectral and time series analysis on quarterly (1950- 1984) and annual data (1871- 1984) in Canada found no empirical evidence to show that government spending had an effect on long term interest rate.

Giffin et al (1981) analyzing time series data over the years 1959-1979 in the US, reported that there was no significant correlation between deficit spending and inflation rate.

Ball and Markiw (1995) presented empirical evidence which showed that large deficits over the period 1982 to 1994 was accompanied by a decline in investment, export, and private saving..

Eisner (1989) and Domar (1993) argued that deficit spending can improve the economy at a time of economic slowdown

Hoelscher (1986), using regression analyses on time series data over the period 1953-1984 in the US, reported that deficit caused long term interest rate to rise. The author was of the opinion that other factors like short term interest rate and inflation were additional factors that may have affected long term interest rate.

Keith (2005 examined the link between deficit and inflation and came to the conclusion that there is little to no link between the two for the US economy.

Hutchison and Pyle (1984), Ford and Lawton (1995), and Tanzi and Fanizza (1995) provided empirical evidence indicating that higher deficits in industrial countries have increased interest rates.

Studies by Barro and Sali-I-Martin (1990) and Evans (1987) support the Ricardian view, namely that a tax induced budget deficit has no effect on interest rate.

Using Granger causality analysis on US data between 1947 and 2002, Liu et al (2008) reported that public expenditure had a positive effect on the GDP. However, GDP did not have any effect on increasing the public expenditure.

Barth and Wells (1999) argued that budget deficit increases interest rate, which can lead in turn to a reduction in investment. This can have a negative effect on economic growth and exports. Barrow (1974) argued that bond-financed deficits will have no effect on economic investment or export.

Palley (2011) argued that deficit financed public investment is needed for economic growth and that austerity measures slows growth. Taylor et al. (2012) presented evidence showing that an increase in public spending in the US had a positive effect on economic growth.

DATA

Data for the United States on federal spending relative to revenue (spending – revenue) and GDP was in billions of dollars. Positive values for spending indicated deficit spending and negative values non-deficit or surplus spending. The data were obtained from the on line source ftp://ftp.bls.gov/pub/special.requests/cpi/cpiai.txt

http://www.economagic.com/em-cgi/data.exe/blsln/lns14000000

http://www.usgovernmentspending.com/spending_chart_1930_2016USk_13s1li011mcn_G0f

Data was over the years 1930-2010. Spending was mostly deficit spending (DS). Plots of the GDP, federal deficit spending (DS), unemployment rate (UER), interest rate (INR), and inflation rate (IR) over years are presented in the Appendix.

METHODOLOGY

The SAS software was used in the data analysis. The Johansen cointegration analysis was performed in order to determine if cointegration exists between GDP and deficit spending, unemployment rate, interest rate, and inflation rate. Also, time series transfer function analysis was used to determine the relationship of GDP to federal deficit spending, unemployment rate, inflation (measured as CPI), and one year fixed deposit interest rate. The final time series model included GDP as a function of federal spending and unemployment. Inflation and interest rate had no significant effect on GDP.

Cointegration

Two time series variables that are in a long-run equilibrium relationship are cointegrated. Cointegrated series do not diverge over time. Any divergence is usually short term and eventually the two series come back together. It is important to realize that co-integrated variables may or may not be correlated.

Table 1 presents the co-integration analysis results for GDP and DS using the Johansen cointegration test (Johansen, 1988).

In the US, government deficit spending relative to revenue (DS) and GDP are cointegrated 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 a long-term linear relationship between the two variables. A similar analysis showed also that GDP is cointegrated with each of employment, interest, and inflation rate. However, as is shown from the time series analysis, GDP is influenced

only by DS and UER and not by INT or IR. Note that cointegration does not imply correlation or functional relationship in the short run.

Table 1. Johansen cointegration rank test for government deficit spending (DS) and GDP

Variables	H_0: rank = r	H_a: rank > r	Trace	Critical Value
DS GDP	0	0	86.67	12.21
	1	1	2.80	4.14

Multivariate time series modeling using the transfer function approach

The transfer function analysis is the state of the art modeling approach to determine the functional relationship between series, the input or independent series and the output or dependent series. The interest in this study is in determining if federal spending relative to revenue does have an effect on GDP in the presence of control variables, namely unemployment, inflation, and interest rates. Hence, the input variables are DS, UER, IR, and INR and the output series is GDP. 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 or 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.

The time series analysis is valid only if the series are stationary. The first difference for each of GDP, UER, and INR was stationary as determined by the Dickey-Fuller unit root test and the dampening patterns of the autocorrelation function (ACF) and the partial autocorrelation function (PACF). On the other hand, the second difference for IR was stationary. Therefore, the analysis that follows is based on the first difference for each of GDP, UER, INR and the second difference for IR.

A transfer function model between an output series y and input series x_i (i = 1, 2, …,k) is expressed in general as

$$y_t = \sum_i^k v(B)_i \, x_{it} + a_t \quad (1)$$

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

The function $v(B)_i$ is determined from the cross correlation between xi and y (Wei, 2006).

Once $v(B)_i$ is identified, one can express et in Eq. (1) as

$$a_t = y_t - \sum_i^k v(B)_i \, x_{it} \quad (2)$$

and identify the appropriate model for Eq. (2 from which one can determine the final model in Eq. (1).

RESULTS

First, the transfer function analysis was performed on the full model with GDP(1) as the output or dependent variable and DS(1), UER(1), INR(1), and IR(1.1) as the input or independent variables. Here, GDP(1), DS(1), UER(1), INR(1) indicate the first difference for each series and IR(1,1) the second difference. Results of the analysis showed that INR(1) and IR(1,1) had no significant relationship to GDP(1). Their p values were 0.30 and 0.83, respectively. As a result, INR and IR were deleted from the model and the analysis repeated using GDP(1), DS(1) and UER(1). The resulting model was

$$Gdp(1)_t = - 0.37DS(1)_t - 49.33UER(1)_t + e_t /(1- 0.85B) \text{ ,} \quad (3)$$

where e_t is white noise and B is the backshift operator as explained above.

The model in Eq. (3) satisfied the assumption that the error term is independent of the input series. This was the case since there was no cross correlation between the noise series and the independent or input series.

From the model in Eq. (3), it is seen that deficit spending, DS (1), has a negative effect on GDP(1) (-0.37 with $p< 0.0001$). Also, as expected UER(1) has a negative effect on GDP(1) (-29.33 with $p <0.0001$).

Furthermore, the cross correlation between GDP(1) and DS(1) was -0.404 ($p = .0097$) and that between GDP(1) and UER(1) was -0.542 ($p < 0.0001$). Based on the Granger test (1969), it was found that UER(1) Granger caused GDP(1) and DS(1) Granger caused GDP(1). These results due to the model in equation (1), the cross correlation and the Granger test, are consistent in showing a negative correlation between GDP(1) and each of UER(1) and DS(1). However, it should be noted that correlation does not necessarily mean causation and that is true also of the Granger causation test, in spite of its name.

For forecasting, equation (3) can be expressed as:

$$GDP(1)_t (1-0.85B) = -0.37 DS(1)_t (1-0.85B) – 49.33 UER(1)_t (1-0.85B)_t \quad (4)$$

Or

$$GDP(1)_t = 0.85 GDP(1)_{t-1} -0.37 DS(1)_t + 0.31 DS(1)_{t-1} -49.33 UER(1)_t +41.93UER(1)_{t-1} \quad (5)$$

Also, from the time series analysis, we have that

$$DS(1)_t = 0.212 DS(1)_{t-1} \quad (6)$$

and

$$UER(1)_t = 0.31 UER(1)_{t-1} \quad (7)$$

Hence, from Eqs. (5), (6), and (7), one can predict $GDP(1)_t$ from observations on

$DS(1)_{t-1}$, $UER(1)_{t-1}$ and $GDP(1)_{t-1}$.

From the predicted $GDP(1)_t$, one can obtain GDP_t from the relation

$$GDP(1)_t = GDP_t – GDP_{t-1} \quad (8)$$

CONCLUSION

The Johanssen co-integrated analysis revealed that GDP has a long-term equilibrium relationship with unemployment, interest and inflation rates. This implies that GDP does not diverge over time from unemployment, interest rate or inflation rate. Any divergence is usually short term and eventually the series come back together. This long-run relationship may be due to direct cause and effect or may be due to a third variable or group of variables that were not observed. Likewise, a functional relationship, as represented by Eqs. (5), (6) and (7) may not be due to direct cause and effect.

Of interest is the finding that deficit spending had a significant negative effect on economic growth in the presence of unemployment, interest rate and inflation rate as control variables. The data used (1930-2010) included the great depression and the recent severe recession. It is not clear why deficit spending had a negative effect when interest rate had no effect. It may be that one year interest rate is not long term to show any effect. Krugman (2012) argued that deficit spending did not help the economy in the recent recession because it was not enough to cause an increase in demand and economic growth. In fact data in this study show that since 2002 and especially after 2007 deficit spending grew significantly while the GDP was stagnant or showed weak growth. Based on this, we analyzed the data for the period 1930 to 2006 and for the period 1930 – 2001. In both cases there was no significant relationship between deficit spending and the GDP. Also, there was no significant cross correlation or Granger causation. This indicates that the last years were the contributing factor for the negative relationship between deficit spending and growth. If Krugman's argument is correct, then the observed negative relationship between deficit spending and GDP may not indicate cause and effect. If deficit spending has a negative effect, it would be because of its effect on raising long term interest rate. In a future work we will examine the relationship between government spending and long term interest rate for the time period after 2000 in the US and other countries.

REFERENCES

Ball, Land N. G. Mankiw. 1995. What do budget deficits do?. National Bureau of Economic Research Working Paper 5263.

Barrow, R. J. 1974. Are government bonds net wealth? Journal of Political Economy. 82. 1095-1117.Barro, R. and X. Sali-I-Martin. 1990. World real interest rates. National Bureau of Economic Research Working Paper No. 3317.

Barth, J. R. and J. M. Wells. 1999. Budget Deficits and Debt: A Global Perspective. Edited by Siamack Shojai, Praeger, Westport, Connecticut.

Bernheim, D. B. 1989. A neoclassical perspective on budget deficits. Journal of Economic Perspectives. 3, 55-72.

Cebula . 2008. Determinants of Long-Term Real Interest Rate Yields: The Case of the US. The Icfai Journal of Applied Economics. VII , 37-49.

Collins, B. 1999. Budget deficits and financial markets. Budget Deficits and Debt: A Global Perspective. Edited by Siamack Shojai, Praeger, Westport, Connecticut.

Domar E. D. 1993. On deficits and debt. American Journal of Economics and Sociology. 52, 475-478.

Eisner, R. 1989. Budget deficits: Rhetoric and reality. Journal of Economic Perspectives. 3, 73-93.

Evans, P. 1985. Do large deficits produce high interest rates? American Economic Review. 7, 68–87.

Evans, P. 1987. Do budget deficits raise nominal interest rates? Evidence from six countries. Journal of Monetary Economics.20, 281-300.

Fledstein, S. 1986. Budegt deficits, tax rules and real interest rates . Working Paper Number 1970. Cambridge Mass.: National Bureau of Economic Research.

Ford, R. and D. Laxton. 1995. World public debt and real interest rates. IMF Working Papers WP/95/30. International Monetary Fund, Wahington , D.C.

Giffin P.E, J.H. Macomber, and R.E. Berry. 1981. An Empirical Examination of Current Inflation and Deficit Spending. Journal of Post Keynesian Economics. IV, 63-67.

Granger, C.W.J. 1969. Investigating Causal Relations by Econometric Methods and Cross-Spectral Methods. Econometrica, 34, 424-438

Hoelscher, G. 1986. New evidence on deficits and interest rates. Journal of Money Credit and Banking. 18, 1–17.

Hutchison, M. M, and D. H. Pyle. 1984. The real interest/budget /deficit Link: International evidence., 1973-1982. Economic Review (Federal Reserve Bank of San Francisco), No 4.

Johansen, S. 1988. Statistical analysis of cointegration vectors. Journal of Economic Dynamics and Control. 12, 231-254.

Kiani, K.M. 2009. Federal budget deficits and long-term interest rates in USA. The Quaterly Review of Economics and Finance. 49, 74-84

Keith Sill. 2005. Do Budget Deficits Cause Inflation? Business Review Q3, 26-33.

Krueger A. B. 2003. Big Federal Deficits, Bigger Risks. New York Times, March 6.

Krugman, P. 2012. End This Depression Now. W.W. Norton and Company, New York, London.

Liu, L. Chih-hung, C. E. Hsu, and M. Z. Younis. 2008. The association between government expenditure and economic growth: Granger causality test of US Data, 1947-2002. Journal of Public Budgeting, Accounting, and Financial Management. 20, 439-452.

Makin, J. 1983. Real interest rates, money surprises, anticipated inflation, and fiscal deficits. Review of Economics and Statistics. 65, 374–384.

MacAvoy, P. 2003. Don't Just Stand There. Treasury Secretary William E Simonand Fiscal Policy During Stagflation 1975-76. Atlantic Economic Journal. 31, 213-218.

Palley, Thomas. 2011. Budget deficit alarmism in sabotaging growth. Challange, Vol. 54, 6-31.

Plosser, C. 1982. Government financing decisions and asset returns. Journal of Monetary Economics. 9, 325–352.

Pollin, R. 2012. US government deficits and debt amid the great recession: what the evidence shows. Cambridge Journal of Economics. 36, 161–187.

Posner M. V.1987. A Survey of the Debate, in Boskin Michael J, Flemming John S and Gorini Stefano (Eds.), Private Savings and Public Debt, pp. 395-414, Basil Blackwell, Oxford, UK.

Siklos, P.L. 1988. The deficit-interest rate link: empirical evidence for Canada. Applied Economics. 20, 1563-1577. Tanzi, V. and D. Fanniza. 1995. Fiscal deficits and public debt in industrial countries, 1970-1994. IMF Working Papers WP/95/30. International Monetary Fund, Wahington, D.C.

Taylor, L. C. R. Proano, L. de Carvalho, and N. Barbosa. 2012. Fiscal deficits, economic growth and government debt in the USA. Cambridge Journal of Economics, 36, 189-204.

Wagner, Adolph . 1892. Grundlegung der politischen Ökonomie. Part 1, vol. 1. 3rd edn. Leipzig: Winter.

Wei, W. (2006). Time Series Analysis: Univariate and Multivariate Methods. Pearson Addison Wesley, New York

APPENDIX

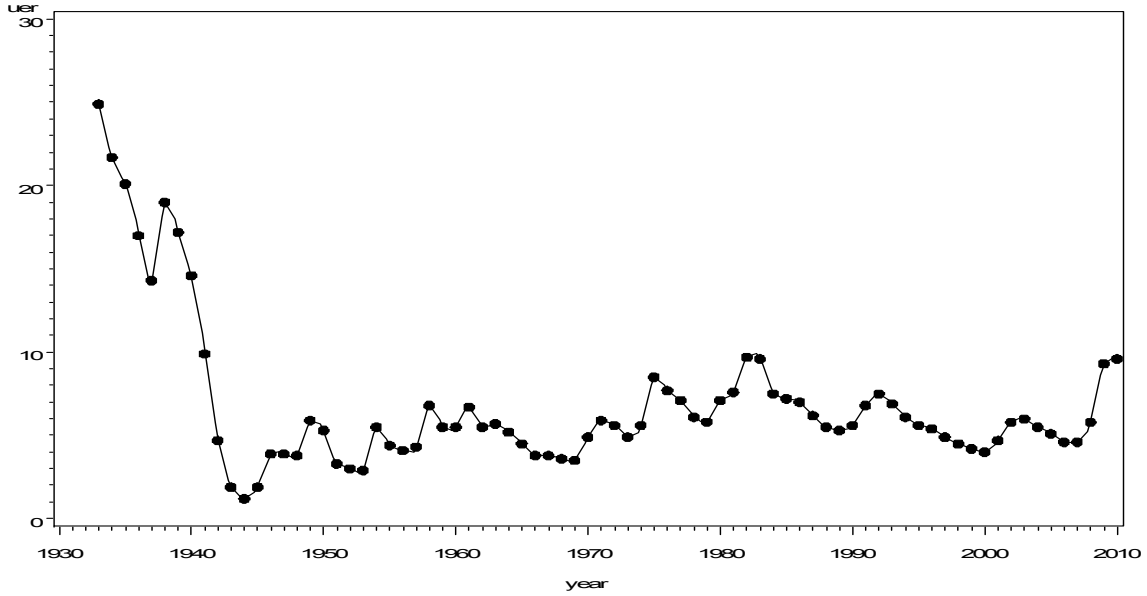


Figure1. Plot of unemployment rate (UER) over years.

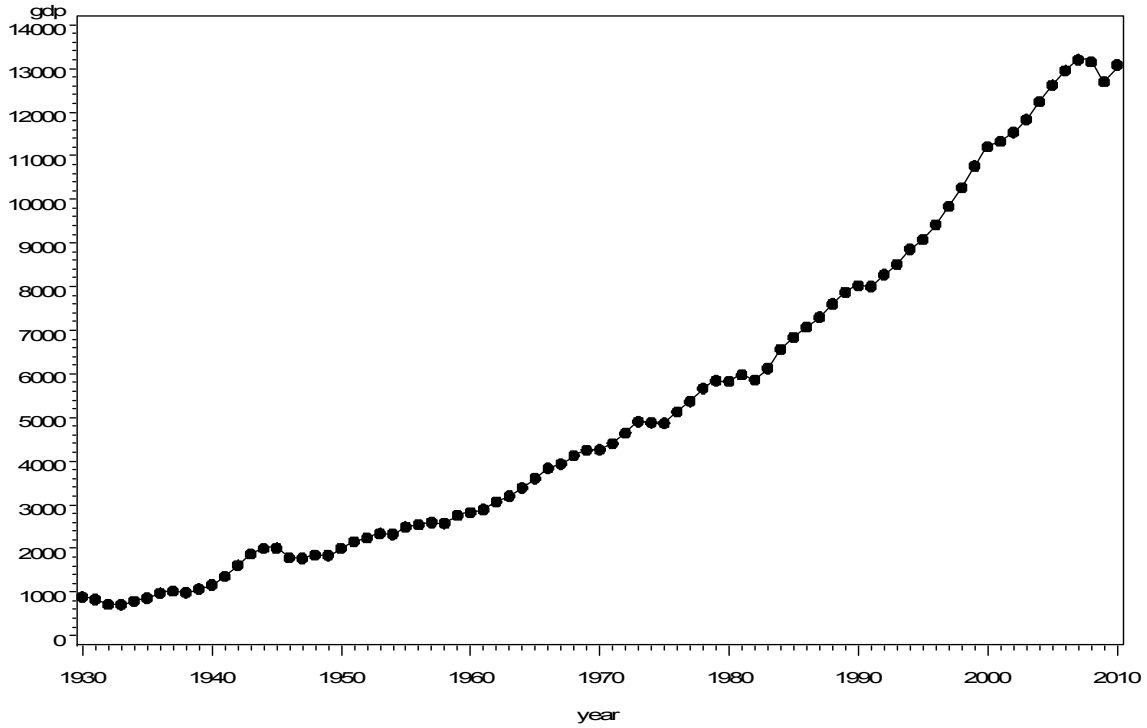


Figure 2. Plot of GDP over years.

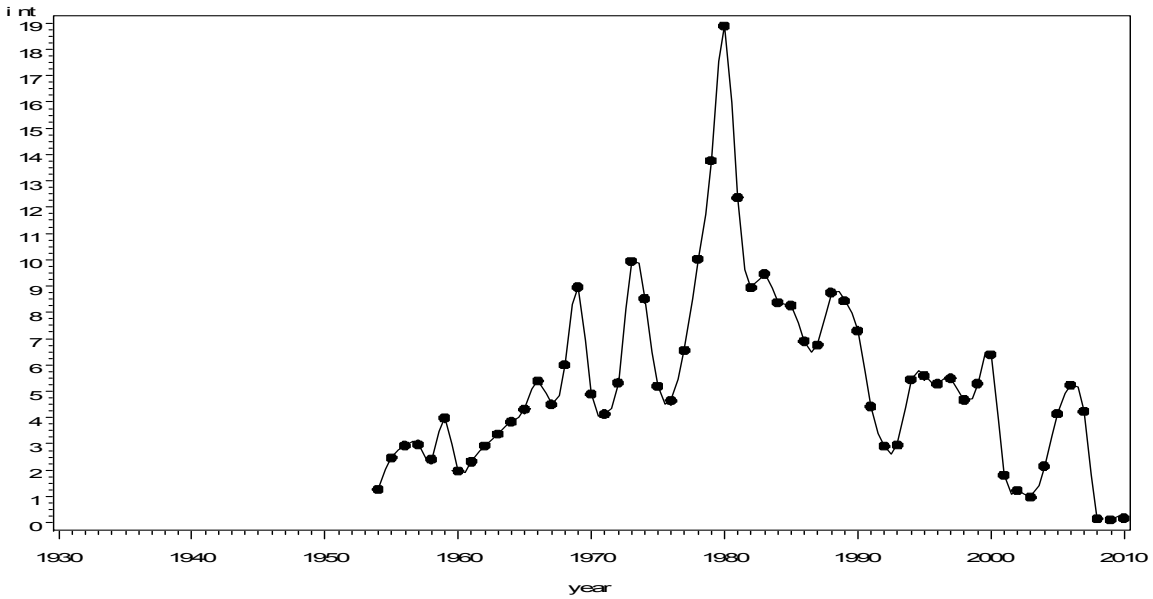


Figure 3. Plot of interest rate (INT) over year.

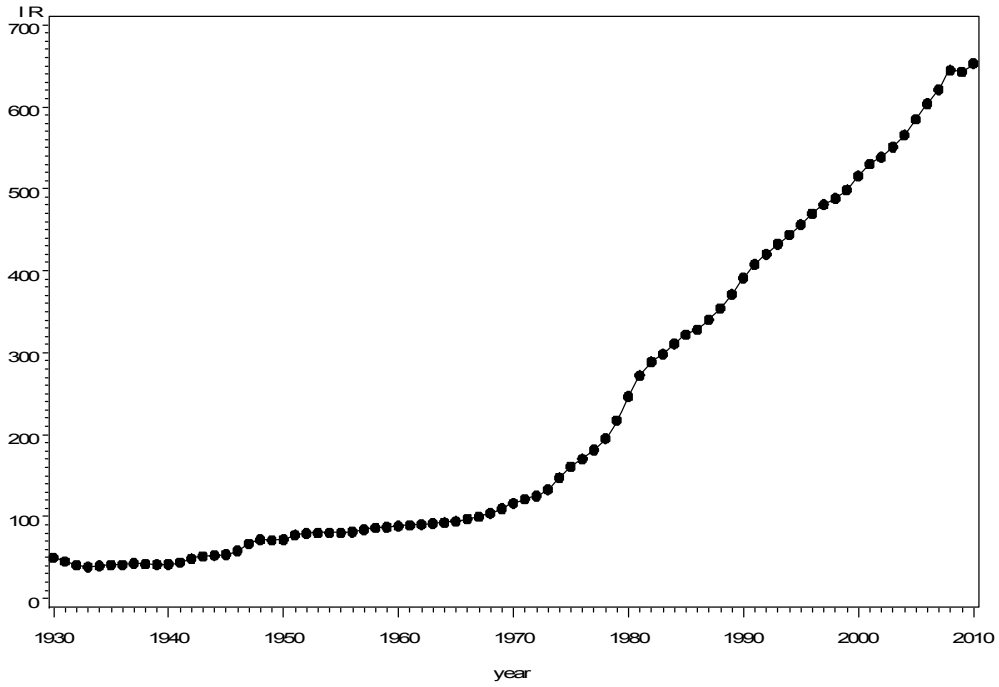


Figure 4. Plot of inflation rate (IR) over years.

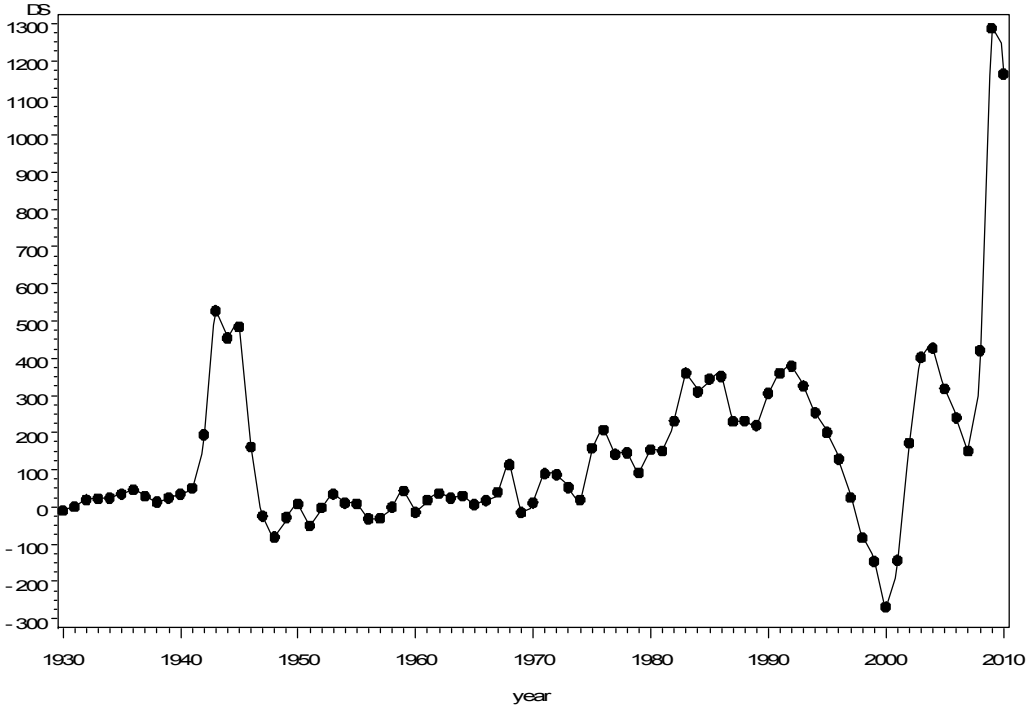


Figure 5. Plot of government deficit spending (spending – revenue), DS, over years.

VARK LEARNING STYLES AND STUDENT PERFORMANCE IN PRINCIPLES OF MICRO- VS. MACRO-ECONOMICS

Ambrose Leung, Mount Royal University
Maureen McGregor, Mount Royal University
David Sabiston, Mount Royal University
Stefano Vriliotis, Mount Royal University

ABSTRACT

Principles of economics courses are known to be taught primarily using a lecture-based format with strong emphasis in the presentation of visual materials such as graphs and tables. While students with certain learning styles can likely appreciate this unique style of presentation, others may find it difficult to comprehend and become frustrated. Evidence has shown that a mismatch between the method used to present course materials and a student's learning style can adversely affect the student's performance in a principles of economics course. However, the literature has not distinguished the potential difference of student performance in principles of micro- versus macro-economics and its relationship with student learning styles.

Using a sample of students from principles of economics courses taught at Mount Royal University in Calgary, Alberta, we examine the relationship between student learning styles, using the VARK (visual, aural, reading/writing, and kinesthetic) inventory, and their performance in principles of micro- versus macro-economics courses. The purpose of this study is to identify whether different student learning styles are related to the performance of students in principles of micro- versus macro-economics.

JEL classification: A10, A22.
Keywords: Learning styles; teaching economics

INTRODUCTION

Students who have taken both courses in principles of microeconomics and principles of macroeconomics quite often express their preference for one over the other. Despite the fact that both branches of economics share the common underlying objective of allocating scarce resources to their best possible uses, microeconomic and macroeconomic analysis tend to require different approaches in terms of information processing that can be associated with differences in personality and learning styles (Bisping and Eells, 2006). The study of microeconomics, which is about how individuals and businesses make decisions, tends to follow logical sequences that are highly structured. The study of macroeconomics, which is about understanding aggregate behaviour of the economy, tends to encourage debates over diverse viewpoints generated by different schools of thoughts. Such differences in the nature between microeconomics and macroeconomics are clearly shown in the textbooks written for the two fields, in that there is much

more variation in the content across macroeconomics textbooks than that in microeconomics textbooks. While the content of microeconomics textbooks appear to be quite standardized using similar logical sequence, macroeconomics textbooks show much greater variety in content, depending on the different viewpoints adopted by the authors. The differences in the content and approaches used in microeconomics and macroeconomics raise the question that students with different personality types, thinking and learning styles may show different preferences for the two fields of study. The literature has seen some studies using personality types to explain such differences. The goal of this paper is to further examine whether preferred learning styles by students affect their course performance in principles of microeconomics versus macroeconomics.

The paper is organized as follows. Section 2 reviews the literature and provides the motivation for the empirical analysis in this paper. Section 3 contains results from the empirical analysis. Section 4 provides interpretation of the empirical results and suggestions for further research.

LITERATURE REVIEW

A typical introductory economics course is lecture oriented (Becker, 2000; Lopus and Hoff, 2009) and relies heavily on visual presentation of information with graphical analysis (Boatman, Courtney, and Lee, 2008; Fleming, 1995). Yet students often show preference for either microeconomics or macroeconomics. Research suggests that course preference and performance of students can vary in relations to their thinking and learning styles. Zhang (2004), for example, found that students with specific thinking styles preferred specific teaching styles. According to Zhang, students with a creative thinking style tend to prefer a learning-oriented and student-focused teaching style. Students with a conformity thinking style tend to prefer a teacher-focused teaching style that emphasizes the transmission of information.

Research in economic education has examined the relationship between learning styles of students and teaching styles of instructors in principles of economics courses. A study by Charkins, O'Toole, and Wetzel (1985) identified three types of student learning styles (dependent, collaborative, and independent) and found that the larger the gap between an instructor's teaching style and a student's learning style, the worse the student's performance in the introductory economics course.

Other studies subsequently examined the relationship between personality types and student performance in principles of economics courses. A common feature among these studies is the use of the Myer-Briggs Type Indicator (MBTI) to identify personality types of individuals. Borg and Shapiro (1996) used a sample of 119 students and found that gender is not a significant factor to determine student performance in principles of macroeconomics once MBTI personality types are accounted for. Ziegert (2000) reached the same conclusion from using a sample of 617 students in principles of microeconomics. McCarty, Padgham, and Bennett (2006), however, found that matching gender of students and instructors significantly improved student performance in both principles of microeconomics and macroeconomics after accounting for MBTI personality types. Another common finding is that personality types were found to be related to student performance. For example, introverts were found to perform better in principles of economics than extroverts (Borg and Shapiro, 1996; Ziegert, 2000). However, it has also been suggested that

personality traits are related to student performance in principles of macroeconomics, but not microeconomics (Bisping and Eells, 2006).

Although certain relationships have been detected between personality types and course performance, it is possible that such relationships exist because MBTI is an indirect indicator of student learning styles through an evaluation of personality. More direct measure of learning styles may be useful in explaining this kind of relationship.

Fleming and Mills (1992) developed an inventory of learning styles know as VARK. The four modes in VARK are visual (V), aural (A), read/write (R), and kinesthetic (K). These modes are frequently referred to as a person's "sensory modality preferences." A person may show no preference, unimodal, or multiple modes of sensory preferences. According to Fleming (1995), students with a visual preference learn best from presentation of materials using graphs, charts and diagrams; aural learners prefer to receive information through listening; read/write learners prefer to take in information through writing and reading from printed words; kinesthetic learners gain better understanding of materials through concrete examples and applications. The most recent version of the VARK questionnaire consists of 16 questions and identifies a person's preferred method or mode of presenting and processing information.

The VARK questionnaire has been widely applied to explore issues related to learning style of students. Some studies showed no gender difference in the numbers or types of sensory differences (Bhaskar, 2011; Slater, Lujan, and DiCarlo, 2007), while others found gender differences in learning style preferences (Dobson, 2010; Rogers, 2009). Attempts have also been made to identify the relationship between VARK learning style preferences and student performance in university courses. For example, Dobson (2010) found that a strong kinesthetic learning style had a significant negative relationship with performance in physiology courses among a sample of 64 students; but Eudoxie (2011) found no significant relationship between VARK learning style preferences and course performance among a sample of 62 students studying soil management science. Other studies used the VARK inventory to show that understanding students' learning style preferences can help to improve the communication of course materials and the educational experience of students. (Dobson, 2010; Rogers, 2009)

Boatman, Courtney, and Lee (2008) used the VARK inventory of learning styles developed by Fleming and Mills (1992) to assess the relationship between student learning styles and their performance among 211 students from a mix of introductory microeconomics and introductory macroeconomics courses. They conclude that students who are visual learners perform better in introductory economics courses and suggest that this result is partly due to the fact that a significant portion of the concepts are taught using a graphical analysis approach. Another observation made by the authors is that once students' learning styles have been addressed, there appears to be no gender-based differences in student performance in introductory economics. This is an interesting point because such finding seems to be consistent with the suggestions from earlier literature in that gender has been found to have no significant relationship with performance in principles of economics courses once personality types are accounted for (Borg and Shapiro, 1996; Ziegert, 2000).

In short, the existing literature has come up with two main findings about the relationship between performance in principles of economics courses and personality types/learning styles. 1) Personality types are related to student performance, and the relationship may be different between

microeconomics and macroeconomics. 2) Gender appears to have no significant relationship with student performance once personality types or learning styles have been taken into account. However, it is not clear whether the relationship between learning style preferences and student performance is different for microeconomics and macroeconomics, which is the main issue to be examined in this paper.

METHOD AND DATA

Our study aims to further examine the relationship between VARK-based student learning styles and student performance in first year university microeconomics versus marcroeconomics courses.

The data in this study came from student surveys based on version 7.1 of the VARK questionnaire developed by Fleming and Mills (1992). The questionnaire consists of sixteen questions that identify the preferred learning styles of students. The data was collected from students enrolled in principles of microeconomics and macroeconomics courses at Mount Royal University in Canada. 1472 students participated in the survey over an 18 months period. Participants were requested to respond to the VARK questionnaire along with information on their age and gender. Each of the sixteen multiple choice questions on the VARK questionnaire has four possible choices that imply preferences for visual, aural, reading/writing, and kinesthetic learning styles respectively. Participants were instructed to choose all the answers that apply to them and not be limited to just one answer to each question. Hence the raw score on each of the sensory modality (i.e., V, A, R, and K) can range from 0 to 16 for each participant. Participants' final grades, measured as a percentage, were collected from the instructors at the end of the term.

Table 1 provides the summary statistics of the variables. Observations used in the empirical analysis must satisfy two criteria. First, observations with any missing variable were eliminated. Second, only those students who passed the course (50% and above) were included. Students who failed or withdrew from the course did so for many different possible reasons that are beyond our control; these observations were therefore eliminated to minimize potential bias.

Table 1: Summary Statistics of Variables

Variables	Mean	Standard deviation	Minimum	Maximum
Age	20.71	3.40	17	46
Gender (1 = female)	0.40	0.49	0	1
Visual (V)	4.92	2.91	0	16
Aural (A)	5.57	2.90	0	15
Reading/writing (R)	5.37	2.92	0	15
Kinesthetics (K)	6.43	2.77	0	16

The final sample used for the analysis consists of a total of 910 participants from first-year economic courses, of which 645 were from microeconomics and 265 were from macroeconomics, with an average age of 20.71 years old. There were slightly more male (60%) than female participants. The raw scores from the four sensory modalities show that participants are more

likely to show a preference for the kinesthetic learning style (mean = 6.43), and least likely to show a preference for the visual learning style (mean = 4.92).

RESULTS

The purpose of the regression analysis is to examine the relationship between total percentage grade and learning style preferences of students. According to Fleming (1995), a strong preference for a learning style can be identified by a score obtained on a learning style that is at least four points above the score of any other learning style. The raw scores on V, A, R, and K were therefore recoded according to Fleming's suggestion described above in our regression analysis. Each of the V, A, R, and K variables was recoded such that a value of 1 indicates a strong preference for a specific learning style, a value of 0 was recorded otherwise. We divided the sample into those students who took microeconomics from macroeconomics. The results are summarized in Table 2.

From the sample of students who passed principles of microeconomics, the results from Table 2 show that age and gender have a significant positive relationship with student performance in the course. Each year of increase in age tends to raise total percentage grade by half percentage point. Females on average receive 2.26 percentage points higher than males. The results for microeconomics show that none of the learning style preferences shows statistically significant relationship with total percentage grade. The adjusted R^2 shows that the explanatory variables included in the analysis account for about 3% of the variation in the final percentage grade from principles of microeconomics.

Table 2: Relationship between Total Percentage Grade and Learning Style Preferences

Variables	Microeconomics Coefficients (t-statistics)	Macroeconomics Coefficients (t-statistics)
Age	0.43*** (3.45)	0.30 (1.81)
Gender (1 = female)	2.26*** (2.46)	0.30 (0.23)
Visual (V)	-0.25 (-1.43)	-1.43 (-0.81)
Aural (A)	0.20 (1.16)	1.37 (0.87)
Reading/writing (R)	-0.02 (-0.10)	-0.53 (-0.32)
Kinesthetics (K)	0.19 (0.98)	3.03** (2.19)
Constant	62.38*** (22.53)	64.66*** (17.52)
Adjusted R^2	0.029	0.034
Number of observations	645	265

**significant at the 5% level
***significant at the1% level

From the sample of students who passed principles of macroeconomics, only the kinesthetic learning style shows a significant positive relationship with total percentage grade. The empirical results imply that students with a strong preference for the kinesthetic learning style significantly increase the total percentage grade by about 3 percentage points. None of the other learning style preferences as well as age and gender show statistically significant relationship with performance in principles of macroeconomics. The adjusted R^2 shows that the explanatory variables included in the analysis account for about 3.4% of the variation in the final percentage grade from principles of macroeconomics.

The findings here show that different factors are related to student performance in microeconomics versus macroeconomics. Age and gender are positively related to student performance in principles of microeconomics. A strong preference for the kinesthetic learning style is positively related to student performance in principles of macroeconomics.

DISCUSSION

The goal of this paper is to analyze the relationship between student learning styles and student performance in principles of economics courses, and to examine potential differences in such relationship between principles of microeconomics and macroeconomics. First year students in both microeconomics and macroeconomics classes, at Mount Royal University, were requested to complete on a voluntary basis the VARK (visual, aural, reading/writing and kinesthetic) survey, along with information on their age and gender. The learning style preferences of each student were recorded, along with the final grade they achieved in the course.

Regression analysis was used with a sample of 910 students to determine whether the different learning styles as well as the age and gender of students are related to their performance in first year economic courses. Students who passed principles of microeconomics and principles of macroeconomics were analyzed separately.

The results from the microeconomics sample, with 645 students, show that age and gender have statistically significant relationship with final grades, while none of the factors representing the four different learning style preferences achieved statistical significance. Hence findings from the microeconomics sample in this study appears to provide different implications than earlier studies that suggested gender as an insignificant factor in relations to performance in introductory economic courses once learning styles or personality types were taken into account (Borg and Shapiro, 1996; Ziegert, 2000; Boatman et al., 2008). Furthermore, the results from the microeconomics sample show that females achieved significantly higher grades than males, which is different from the suggestions of some studies that economics has been a male-dominated discipline that favours male students (McCarty, Padgham, and Bennett, 2006).

The results from the macroeconomics sample, with 265 students, show that a preference for the kinesthetic learning style is the only factor that has a significant relationship with final grade. None of age, gender, or any of the other learning styles achieved statistically significant results. Hence our findings are different from the suggestion by Boatman et al. (2008) that students with strong visual preference performed better in introductory economics courses.

The empirical analysis in this paper shows some interesting differences between microeconomics and macroeconomics regarding factors that are related to student performance in

principles of economics courses. In examining the relationship between personality types and student performance, Bisping and Eells (2006) found that personality traits had no significant relationship with performance in principles of microeconomics, but personality traits appeared to play a role in determining performance in principles of macroeconomics. In particular, the authors suggested that students who performed well in macroeconomics had personality traits that favour real and tangible information rather than general patterns. This description shares some similarities with Fleming's (1995) description of the kinesthetic learners who learn concepts and theories through applications and real life examples. Hence our finding that kinesthetic learners performed well in macroeconomics is consistent with evidence from existing literature regarding personality traits.

Why do students with personality traits and learning styles that favour tangible real life applications tend to perform better in principles of macroeconomics, but not microeconomics course? One possible explanation is that there tends to be more coverage of macroeconomic topics and policies in the news media such as changes in interest rate, unemployment, inflation, and so on. Perhaps such abundance of available information makes it easy for instructors to access and discuss "real life issues" in class. More research in this area will be useful to further explore such relationship.

Why do age and gender matter in relations to performance in principles of microeconomics, but not macroeconomics? Perhaps most students take microeconomics before macroeconomics, and older students are more adaptable to understand economics, given their maturity level and stronger work ethics. The same phenomenon may apply along the gender line in that young adult females are more mature than males in the same age group in dealing with initial exposure to the study of economics. This is another interesting issue that requires further research.

More research effort needs to account for the different learning environments offered in first year economics courses as well as instructor characteristics in terms of gender and learning styles. This should provide a more comprehensive picture of what is required to encourage more productive learning, given the different learning styles of our students.

REFERENCES

Becker, W. E. (2000). Teaching economics in the 21st Century. *Journal of Economic Perspectives,* 14(1), 109-119.

Bhaskar, S. (2011). Implications of perceptual learning style preferences on management pedagogy. *International Journal of Research in Commerce & Management,* 2(2), 109-114.

Bisping, Timothy O., and James B. Eells (2006). Personality type as a determinant of student performance in introductory economics: Macroeconomics vs. microeconomics. *Journal of Economics and Economic Education Research*, 7(1), 3-28.

Boatman, K., R. Courtney, and W. Lee (2008). "See how they learn": The impact of faculty and student learning styles on student performance in introductory economics. *The American Economist,* 52(1), 39-48.

Borg, M. O., and S. L. Shapiro (1996). Personality type and student performance in principles of economics. *Journal of Economic Education,* 27(1), 3-25.

Charkins, R. J., D. M. O'Toole, and J. N. Wetzel (1985). Linking teacher and student learning styles with student achievement and attitudes. *Journal of Economic Education,* 16(2), 111-120.

Dobson, J. L. (2010). A comparison between learning style preferences and sex, status, and course performance. *Advances in Physiology Education,* 34(4), 197-204.

Euodxie, G. D. (2011). Learning styles among students in advanced soil management class: Impact on students' performance. *Journal of Natural Resources and Life Sciences Education,* 40(1), 137-143.

Fleming, N. D. (1995). I'm different; not dumb: Modes of presentation (VARK) in the tertiary classroom, In A. Zelmer (Ed.), *Research and Development in Higher Education, Proceedings of the 1995 Annual Conference of the Higher Education and research Development Society of Australasia (HERDSA), HERDSA* 18 (pp. 308-313).

Fleming, N. D., and C. Mills (1992). Not another inventory, rather a catalyst for reflection. *To Improve the Academy,* 11, 137-155.

Lopus, J., and J. Hoff (2009). Is there a disconnect between teaching styles and learning styles in high school economics classes? *The Journal of Private Enterprise,* 24(2), 157-164.

McCarty, C., G. Padgham, and D. Bennett (2006). Determinants of student achievement in principles of economics. *Journal for Economics Educators,* 6(2), 1-9.

Rogers, M.A. (2009). A preliminary investigation and analysis of student learning style preferences in further and higher education. *Journal of Further and Higher Education,* 33(1), 13-21.

Slater, J. A., H. L. Lujan, and S. E. DiCarlo (2007). Does gender influence learning style preferences of first-year medical students? *Advances in Physiology Education,* 31(4), 336-342.

Zhang, L-F. (2004). Do university students' thinking styles matter in their preferred teaching approaches? *Personality and Individual Differences,* 37(8), 1551-1564.

Ziegert, A.L., 2000. The role of personality temperament and student learning in principles of economics: Further evidence. *Journal of Economic Education* 31(4): 307-322.

AN EXPLORATORY STUDY OF DIFFERENCES IN STUDENTS' VIEWS OF THE MARKET SYSTEM

John G. Marcis, Coastal Carolina University
Alan B. Deck, Bellarmine University
Daniel L. Bauer, Bellarmine University
Vicki King-Skinner, Coastal Carolina University

ABSTRACT

This study examines data gathered in an anonymous in-class survey of first-year university students regarding the perception of markets. The study employs a survey instrument developed by Lephardt and Breeden (2005). The purpose of the study is two-fold. First, the study investigated the existence of differences in the perception of the role of markets along gender lines. This study found that although male students generally had a more favorable impression of markets, only token statistical evidence was found for the existence of these differences. Second, the study investigated the existence of differences in the perception of the role of markets by major field of study. This study found the existence of differences in the perception of markets to be somewhat striking.

THE SURVEY INSTRUMENT AND ASSOCIATED MATERIAL

This study uses the survey instrument (Market Attitude Inventory, MAI) developed by Breeden and Lephardt (2002) and Lephardt and Breeden (2005). The 2002 paper was an empirical study that used a survey instrument first developed by the authors in 1992 and refined over the next decade. The 2002 study involved 406 student responses in three different economics courses during two time periods (1992 and 1999). The authors found "significant differences in attitudes between demographic subcategories and between classes of students, as well as changes in attitudes over the time elapsed" (Breeden and Lephardt 2002, 154). The 2005 study provides the development and underpinnings of the survey instrument used in the 2002 paper. In many ways, the methodology set forth in the Lephardt and Breeden study of 2005 preceded the 2002 study. The authors noted two factors that motivated them to develop the MAI. One factor was a long-term research agenda involving "the evaluation of the relationship between an individual's attitudes toward the market system and achievement of economic success within that system" (Lephardt and Breeden 2005, 63). The second factor was the absence of any valid survey instrument "that measured the values and attitudes people hold toward the market system" (Lephardt and Breeden 2005, 63).

The survey instrument has two sections. The first section of the survey requested demographic data from the individual respondent. Specific questions pertained to the respondent's gender, age, ethnicity, and major field of study.

The second section of the survey instrument was a slightly modified version of the MAI developed by Lephardt and Breeden (2005). The original MAI had 22 statements that measured attitudes towards the market system. For each of the 22 statements, students were asked on the survey instrument to "indicate your level of agreements to each statement by writing a number between

'0%' and '100%'' for the statement, with '0' indicating "strongly disagree" with the statement and "100%' strongly agree with the statement. Some of the statements portray a positive slant towards a market economy and some portray a negative slant towards a market economy (Thomas and Campbell 2006, 33). The 22 statements in the MAI are prefaced with the clause "In my opinion, the market system in the U.S. …" (Lephardt and Breeden 2005, 68). Breeden and Lephardt found "students in more advanced business classes having the most pro-market attitudes" (Breeden and Lephardt 2002, 169). However, since this study focused on first-year students, and first-year students, with minimal exposure to either business or economics in the high school curriculum, might focus on the word "market" in the introductory clause and lose sight of the study's emphasis. Consequently, it was decided that the introductory clause be massaged to read "In my opinion, the *economic system* in the United States:" (emphasis added). The wording of the 22 statements, however, did not change from the original MAI. Five additional statements that relate to the role of the federal government in a market-based economy were introduced. Hence, the wording of the introductory clause to these five statements (23-27, inclusive) was revised to read "In my opinion, the federal government of the United States should:" All 27 statements are listed in Table 1.

Table 1
The MAI Survey Questions

Situation/Scenario

In my opinion, the economic system in the United States:

1. … leads to an unfair distribution of income.
2. … rewards people fairly for their productivity and hard work.
3. … encourages unethical business behavior.
4. … leads to quality and technological advancement in products and services.
5. … leads to inadequate amounts of important public services (like police, roads, fire prevention).
6. … provides opportunities and incentives for success.
7. … encourages greed and excessive materialism.
8. … allows equal access to work opportunities.
9. … leads to erratic cycles of growth and then decline in economic activity.
10. … raises the living standard for most people.
11. … leads to monopoly power among businesses.
12. … leads to an efficient use of resources.
13. … encourages the abuse of he environment.
14. … leads to unemployment and worker insecurity.
15. … leads to excessive risk of business failure.
16. … requires a lot of government control to work well.
17. … allows too much foreign competition.
18. … provides consumers the goods and services they want.
19. … provides employment opportunities for all who desire work.
20. … encourages innovation and the development of new businesses.
21. … provide goods and services at an affordable price.
22. "Overall, I believe that the economic system in the United States is a fair and ethical system."

In my opinion, the federal government of the U.S. should:

23. … take a greater responsibility for solving the problems in society.
24. … balance the budget every year.
25. … limit the importation of foreign products to protect the American economy.
26. … use tax policies to promote a more equal distribution of income.
27. … provide a job to any one who wants one.

SURVEY RESULTS

The expanded version of the MAI was administered anonymously during the second week of the semester to students in eight sections of a freshman-level course. The sections ranged in size from 18 to 25 students.

A total of 186 survey instruments were returned but one survey was discarded for incomplete responses. Of the 185 viable surveys, 84 were from young women and 101 returned from young men. Approximately 90 percent (164) of the respondents self-identified themselves as Caucasian while seven respondents self-identified themselves as African-American and seven more self-reported themselves as Hispanic (or Latino/Latina). Over 55 percent of the students (102) indicated they were planning to major in an area within the College of Business.

Table 2 provides the 27 survey statements. The table is arranged in such a manner as to provide sample characteristics (mean and standard deviation) for each statement for five different cohorts (specifically, Overall, Females, Males, Business, and Non-Business). The cohorts are arranged in such a manner that it allows a test of the differences in the mean responses for both gender and major field of study (specifically, Business and Non-Business).

Examining Differences in Mean Responses by Gender

Breeden and Lephardt (2002) examined a number of sub-categories among the respondents. One of their findings is that male students tended to be more pro-market than females. King and King (2007) also used the MAI and found "females had less favorable views of free markets than males, although neither group reported particularly strong beliefs" (King and King 2007, 168).

Table 2 allows for the examination of the mean responses along gender lines. In general, for the original 22 MAI statements, the mean responses by males were generally more favorable to the market than were the mean female responses. This was true for statements that held either a positive slant or a negative slant towards the role of markets. However, in only two of the original 22 MAI statements (19 and 21) is the difference in the mean responses *statistically different* at the ten percent level.

This study added five statements (23-27, inclusive) that pertained to the role of the federal government in the United States economy. For all five statements, the mean responses by males were more pro-market than were the mean responses for females. However, for only one of the five statements was the difference between the means found to be statistically significant at the ten percent level.

Examining Differences in Mean Responses by Major

Breeden and Lephardt found one of "the most notable subgroup categories was the Business Major comparison between the 1992 and 1999 sample" (Breeden and Lephardt 2002, 160). Breeden and Lephardt found nine of the differences in means to be statistically significant at the one percent level. Moreover, the authors found that in eight of those instances "the means indicate there is a pervasive decline in positive attitudes towards the market" (Breeden and Lephardt 2002, 160) among Business majors. Table 1 allows for the examination of the mean responses by major college of study (that is, Business and Non-Business). In general, for the original 22 MAI statements, the mean responses by Business majors were generally more favorable to the market than were the mean responses by Non-Business majors. This was true for statements that held either a positive slant or a negative slant towards the role of markets. For six of the original 22 MAI statements was the difference in the mean responses *statistically different* at the ten percent level. One interesting finding is that among the statements with statistically significant outcomes, four had a negative slant towards markets while two had a positive slant towards markets. Statements 1, 3, 7 and 11 have a negative slant towards the role of markets in the functioning of the economy. Specifically, Statement 1 pertained to promoting an unfair distribution of income, Statement 3 commented upon encouraging unethical business behavior, Statement 7 addressed greed and excessive materialism, and Statement 11 referenced the attempts by business to promote monopoly power. The two positive statements pertained to providing employment opportunities for all who desire to work (Statement 19) and encouraging innovation and the development of new businesses (Statement 20). None of the five statements that pertained to the role to the federal government were found to be statistically significant at even the 20 percent level.

The differences in mean responses, including a more favorable view of the role and efficiency of markets by business majors, may be explained by learning experiences found in general management degree programs. AACSB International specifies undergraduate degree programs include experiences in management specific knowledge and skills areas, including domestic and global economic environments of organizations, and, ethical and legal responsibilities in organizations and society. To address these requirements, Business schools typically provide direct, applied knowledge and learning experiences in macroeconomics. In a typical macroeconomics course, students analyze the public sector of the economy, while focusing on the decision making process of government. Included in this experience are analysis and discussions of areas such as the role of government in solving problems - such as market failure, poor information, lack of competition in markets, and economic instability. As a result of these experiences, business majors (relative to non-business majors) should have a more thorough, balanced, and in-depth understanding of the role and limits of the economic system and role of government in the United States.

CONCLUSIONS

The objective of this study was to investigate the existence of differences in the perception of markets along both gender lines and major field of study. This study found male students

generally had a more favorable view of markets than female students but that this difference was not particular strong in a statistical framework. This study also found a pronounced difference in the perception of markets along major fields of study.

REFERENCES

Breeden, Charles H., & Noreen E. Lephardt (2002). Student Attitudes Towards the Market System: An Inquiry and Analysis. *The Journal of Private Enterprise*, 17, 153-171.

King, Amanda S., & John T. King (2007). Attitudes Toward Free Markets and Interest in Pro-Market Organizations: Evidence from Students in Free Enterprise. *The Journal of Private Enterprise,* 23, 165-171.

Lephardt, Noreen E., & Charles H. Breeden (2005). The Market Attitudes Inventory: The Development and Testing of Reliability and Validity. *Journal of Economics and Economic Education Research,* 6, 63-72.

Parker, Darrell, & Martha C. Spears (2002). Personality Type, Gender, and Risk: A Framing Analysis. *Economics and Economic Education Research,* 3, 69-86.

Thomas, M. Kathleen, & Randall C. Campbell (2006). Teacher Training and Market Attitudes in Transitioning Economies. *The American Economist*, 50, 32-41.

Table 2
Response Summaries and Tests of Hypotheses

Situation/Scenario	Cohort	Characteristics: mean	st. dev.	$H_1: \mu_x - \mu_y \neq 0$ Pr > | t |
In my opinion, the economic system in the United States:				
1. …leads to an unfair distribution of income.	**Overall**	**53.45**	**24.56**	
	Females	56.31	22.05	
	Males	51.07	26.34	0.149
	Business	48.75	23.70	
	Non-Bus	59.22	24.51	0.004
2. …rewards people fairly for their productivity and hard work.	**Overall**	**54.95**	**23.94**	
	Females	52.12	22.46	
	Males	57.30	24.97	0.143
	Business	56.18	24.10	
3. …encourages unethical business behavior.	Non-Bus	53.43	23.80	0.440
	Overall	**44.76**	**22.56**	
	Females	46.25	21.23	
	Males	43.51	23.65	0.413
	Business	41.27	23.26	
	Non-Bus	49.04	21.03	0.020
4. …leads to quality and technological advancement in products and services.	**Overall**	**74.34**	**18.14**	
	Females	73.21	19.32	
	Males	75.27	17.13	0.445
	Business	74.36	17.32	
	Non-Bus	74.30	19.20	0.982
5. …leads to inadequate amounts of important public services (like police, roads, fire prevention).	**Overall**	**44.46**	**21.70**	
	Females	44.43	21.82	
	Males	44.49	21.70	0.987
	Business	44.48	21.76	
	Non-Bus	44.43	21.76	0.988
6. …provides opportunities and incentives for success.	**Overall**	**71.25**	**21.58**	
	Females	70.17	20.09	
	Males	72.16	22.81	0.534
	Business	71.21	22.18	

Table 2 (continued)
Response Summaries and Tests of Hypotheses

Situation/Scenario	Cohort	Characteristics: mean	st. dev.	$H_1: \mu_x - \mu_y \neq 0$ Pr > | t |
In my opinion, the economic system in the U.S.:				
9. …leads to erratic cycles of growth and then decline in economic activity.	**Overall**	**59.77**	**20.92**	
	Females	60.12	22.05	
	Males	59.49	20.04	0.838
	Business	59.98	20.61	
	Non-Bus	59.52	21.42	0.882
10. …raises the living standard for most people.	**Overall**	**58.56**	**19.78**	
	Females	58.07	19.32	
	Males	58.96	20.24	0.665
	Business	58.87	18.53	
	Non-Bus	58.17	21.32	0.811
11. …leads to monopoly power among businesses.	**Overall**	**53.91**	**22.92**	
	Females	56.00	23.00	
	Males	52.18	22.83	0.260
	Business	50.25	22.49	
	Non-Bus	58.42	22.78	0.015
12. …leads to an efficient use of resources.	**Overall**	**44.64**	**22.79**	
	Females	45.33	22.21	
	Males	44.06	23.36	0.706
	Business	46.76	22.71	
	Non-Bus	42.01	22.76	0.160
13. …encourages the abuse of the environment.	**Overall**	**53.63**	**24.88**	
	Females	51.13	26.86	
	Males	55.70	23.03	0.214
	Business	51.75	23.68	
	Non-Bus	55.93	26.23	0.258
14. …leads to unemployment and worker insecurity.	**Overall**	**50.03**	**22.27**	
	Females	52.07	23.15	
	Males	48.34	21.49	0.257

Table 2 (continued)
Response Summaries and Tests of Hypotheses

Situation/Scenario	Cohort	Characteristics: mean	st. dev.	$H_1: \mu_x - \mu_y \neq 0$ Pr > | t |
In my opinion, the economic system in the U.S.:				
17. …allows too much foreign competition.	**Overall**	**48.60**	**24.14**	
	Females	51.23	24.31	
	Males	46.42	23.74	0.178
	Business	48.26	24.90	
	Non-Bus	49.01	23.33	0.835
18. …provides consumers the goods and services they want.	**Overall**	**71.68**	**22.65**	
	Females	68.93	25.19	
	Males	73.96	20.15	0.133
	Business	71.18	23.91	
	Non-Bus	72.29	21.19	0.741
19. …provides employment opportunities for all who desire to work.	**Overall**	**50.76**	**27.21**	
	Females	46.26	27.03	
	Males	54.51	26.92	0.040
	Business	54.79	25.88	
	Non-Bus	45.81	28.11	0.025
20. …encourages innovation and the development of new businesses.	**Overall**	**65.24**	**22.08**	
	Females	63.20	22.07	
	Males	54.51	22.07	0.254
	Business	54.79	22.29	
	Non-Bus	45.81	21.53	0.074
21. …provide goods and services at an affordable price.	**Overall**	**54.64**	**21.03**	
	Females	51.19	22.81	
	Males	57.51	19.08	0.041
	Business	56.67	21.68	
	Non-Bus	52.16	20.05	0.147
22. *"Overall, I believe that the economic system in the United States is a fair and ethical system."*	**Overall**	**57.44**	**23.29**	

Table 2 (continued)
Response Summaries and Tests of Hypotheses

Situation/Scenario	Cohort	Characteristics: mean	st. dev.	$H_1: \mu_x - \mu_y \neq 0$ Pr > | t |
In my opinion, the federal government of the U.S. should:				
23. …take a greater responsibility for solving the problems in society.	**Overall**	**59.43**	**27.21**	
	Females	61.21	25.34	
	Males	57.94	28.71	0.409
	Business	58.19	26.33	
	Non-Bus	60.95	28.33	0.493
24. …balance the budget every year.				
	Overall	**72.22**	**23.18**	
	Females	75.36	22.29	
	Males	69.60	23.68	0.093
	Business	72.06	23.20	
	Non-Bus	72.41	23.29	0.919
25. …limit the importation of foreign products to protect the American economy.	**Overall**	**53.92**	**24.43**	
	Females	55.06	23.56	
	Males	52.97	25.20	0.564
	Business	54.51	23.54	
	Non-Bus	53.19	25.60	0.716
26. …use tax policies to promote a more equal distribution of income.	**Overall**	**50.24**	**28.37**	
	Females	50.89	28.71	
	Males	49.69	28.21	0.863
	Business	48.45	25.84	

AN ANALYSIS OF US HOUSEHOLD SOCIOECONOMIC PROFILES BASED ON MARITAL STATUS AND GENDER

Sumaria Mohan-Neill, Roosevelt University, Chicago
Indira Neill Hoch, University of Illinois, Chicago
Meng Li, Roosevelt University, Chicago

ABSTRACT

Previous studies have reported that there are significant differences in income and wealth based on marital status and marital history (Wilmoth & Koso, 2002; Gustman & Juster, 1995; Seigel, 1993; Holden & Kuo, 1996). This paper first examines the separate effects of marital status and gender on the socioeconomic profiles of US households, and then explores the combined effect of the interaction of gender and marital status on those profiles. As expected, the results show that married head of household (HH) families are financially better off than single HH families. However, when gender is introduced, it seems to subtract from the gains of marriage. A married female HH is significantly worse off than married male HH, and is closer to single female HH in income. Furthermore, married female HH has less net worth than single male HH. Single female HH group has lowest income and wealth of all groups and this group constitutes about 22.5% of all households. Single female HH is also the 2nd largest group with children in the US, and due to their relatively low economic resources, these households often struggle with reduced quality of life and educational opportunities. This toxic combination increases the risks of continued inequality and inter-generational propagation of poverty. The most interesting findings are the results of the OLS regression models. They show that while demographic variables such as gender and marital status have significant correlation with wage income, household income and net worth, they are considerably less significant than age and education. Finally, all the demographic correlates (age, education, gender, marital status, etc.) pale in comparison to the influence of wage income and household income on net worth. So, if one is interested in policy prescriptions, substantive analysis should consider all these variables in a real world context, and imagine the scenarios where it would be more effective to enact policy measures for helping with opportunity and inequality.

INTRODUCTION AND SELECTED BACKGROUND LITERATURE

Wealth inequality in the US has been increasing in fits and starts since the 1930s and particularly since the 1980s (Wolff, 1992). Increasing wealth inequality in the United States is driven by the top tail of the distribution becoming increasingly wealthy, resulting in a long tail of those with low or negative wealth (Diaz-Gimenez, Glover & Rios-Rull, 2007). Americans desire less inequality than currently exists (Norton & Ariely, 2011). The social and political implications of wealth inequality extend beyond wealth accumulation (Neill Hoch & Mohan-Neill, 2013).

Wolff (1998) argues that "in a representative democracy, the distribution of power is often related to the distribution of wealth." Likewise, social and political factors that contribute to the unequal distribution of income and wealth are varied and interlocking. For these reasons, it is important to understand the demographic populations that currently show signs of difficulty in accumulating wealth.

Family structure has been shown to correlate with wealth. Married households, with or without children, are wealthier than single households (Diaz-Gimenez et al., 2007). Those married continuously are wealthier than those who have had a marriage dissolve, either by divorce or death (Wilmoth & Koso, 2002). Remarriage mitigates some of the losses accrued by marriage dissolution, but not all. Single individuals who have never married see less wealth accumulation than do those who have been married and reaped the benefits of pro-marriage policy for at least part of their lives. Cohabitation with a partner does not show similar benefits to marriage (Wilmoth & Koso, 2002). Cohabitating, non-married partners (also called Living with Partner or LWP) may share some expenses, but such sharing does not translate into increased wealth accumulation over time. Notably LWP cohabitation does not carry the same tax and policy advantages that marriage does. Women who have never been married see an 86% reduction in their overall wealth, pre-retirement, as compared to men who see a 61% reduction (Wilmoth & Koso, 2002), suggesting a gender bias in wealth accumulation.

Changes in family structure may contribute to increased inequality (Esping-Anderson, 2007; McLanahan, 2004). Single parent households have increased over time, from both never married individuals raising children and marriage dissolution. Single females with children may have increased the number of low income households (McCall & Percheski, 2010). Single females with children see the greatest intragroup income inequality of all family structures (Diaz-Gimenez, Glover & Rios-Rull, 2007). McCall and Percheski (2010) argue that there is "strong support for the hypothesis that increases in single mother families and decreases in married couple families have increased income inequality (p. 337). Wilmoth & Koso's (2002) findings that women, with or without children, see a significant reduction in wealth when they remain unmarried, supports the idea that marriage encourages wealth accumulation. Because LWP has not shown to produce similar benefits, it may be that policy benefits awarded to married couples that are not extended to cohabitating couples encourage wealth accumulation.

The wealthy are more educated (Diaz-Gimenez et al., 2007). However, education alone does not always translate to increased wealth. When considered with age, the young educated tend to have little wealth as they pay back debt acquired during schooling or establishing their households. Wealth being strongly correlated with education should not be taken as a magic bullet for addressing wealth inequality. While education increases earning potential, such education may not translate to financial literacy and increased saving behavior (Lusardi & Mitchell, 2007). In education, as in other variables, contribution to wealth accumulation interlocks with additional variables. There is no straight path to predicting wealth.

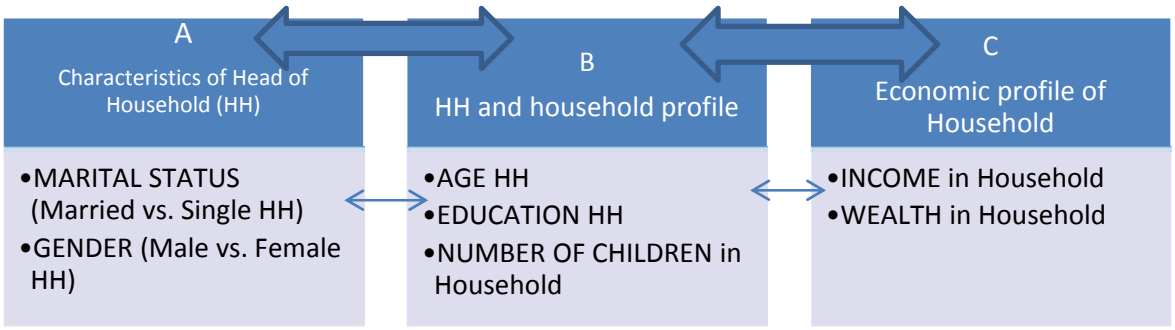
A recent Pew Research study has explored changes in household economics (between 1970 and 2007) as it relates to gender and marital status (Fry & Cohn, 2010). This paper utilizes more current (2010) Federal Reserve's Survey of Consumer Finances (SCF) data to evaluate the profiles of households based on marital status and gender differences. In particular, it will examine the

correlation of those variables with age, education and number of children, income (wage and household), and wealth (net worth).

DATA AND MEDTHODOLOGY

Figure 1 illustrates the variables and relationships which will be analyze and tested in this paper. *Marital status* and gender of head of household (*gender HH*) are the two variables which will be used to compare differences in profiles of US households.

Figure 1. Relationships Analyzed in Model



Research Questions

This paper will address the following specific research questions concerning the demographic and economic status of US households in 2010:

1. What is the impact of *marital status* on household socioeconomic profiles? Using *marital status* as an independent classification variable, it will compare the differences in profiles of *single head of household (HH)* versus *married HH*, with respect to age, education, number of children, income and wealth. It will also test whether the differences in profile variables (e.g., age, education, number of children, HH income, and net worth) of *single HH* versus *married/LWP* households are statistically significant.
2. What is the impact of *gender* on household socioeconomic profiles? Using *gender of HH* as an independent classification variable, it will compare the differences in profiles of *male (HH)* versus *female HH* families, with respect to age, education, number of children, HH income and net worth. It will also test whether the differences in profile variables (e.g., age, education, number of children, HH income, and net worth) of *male HH* versus *female HH* households are statistically significant.
3. What is the impact of *marital status* and *gender* simultaneously on household socioeconomic profiles? It will examine the combined effect of *marital status* and *gender* and compare the differences in profiles of households with respect to age, education, number of children, income and wealth.

4. What are the societal and personal implications, with respect to the current state and trajectory of US household socioeconomic stratification?

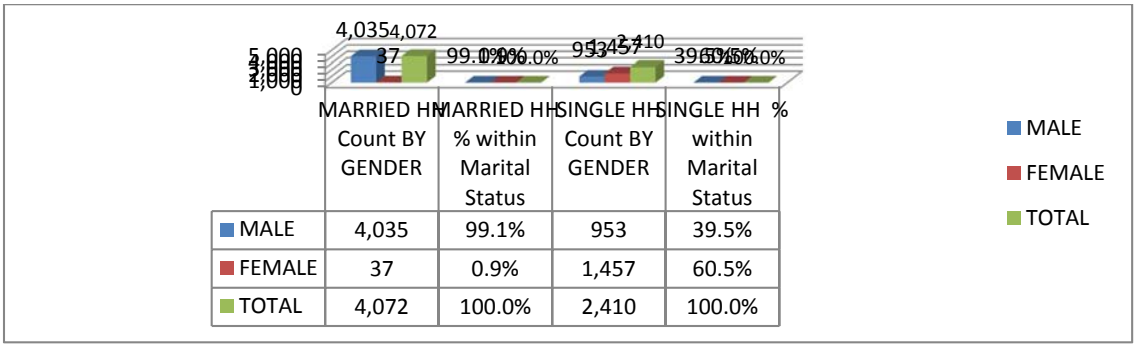
Data and Sample

The data source is the Survey of Consumer Finances (SCF) 2010, which was collected on behalf of the Federal Reserve (http://www.federalreserve.gov/econresdata/scf/scfindex.htm). The SCF survey is conducted every three years (2010 is the latest available dataset).

The sample is comprised of 6,482 households; a little less than two-thirds (62.8%) of SCF sample is classified as "*married HH"*. So, more than one-third of sample US households are classified as "*single HH"* (37.2%). Approximately 23% of all households are headed by a *female HH,* and 77% have a *male HH*. This paper will examine the impact of *marital status* and *gender* separately and then explore the magnitude of their combined effects on household socioeconomic profile and prospects.

Figure 2 illustrates the frequency distributions of household when *marital status* and *gender* are classification variables. The majority of *married HH* are headed by *male HH* (99.1%) and less than 1% is *female HH*. There are more *female HH* in *single HH* (about 60%) compared to about 40% *male HH* in the *single HH marital status* category.

Figure 2. Frequency Distribution of Households by *Marital Status* and *Gender*



The following variables are examined in this study:

1. *Marital status*: (a) Single, not married nor LW P (living with partner, (b) Married, married or LWP (living with partner).
2. *Gender* (Head of household): (a)Male, (b) Female
3. *Age* (HH): reported in years and also in age categories
4. *Education* (HH): reported in years and also in categories
5. *Number children inclusive in household.*
6. *Income:* (last 12 months household income in dollars) which is the total amount of wages (i.e. wage income), investments, transfers etc.
7. *Wage Income* (HH in dollars): included in *Income* (6)
8. *Net worth* (wealth in dollars): All assets minus all debts

Both mean and median values of central tendency will be reported in this paper. However, due to the large variance in some economic variables, median values are the preferred measure to

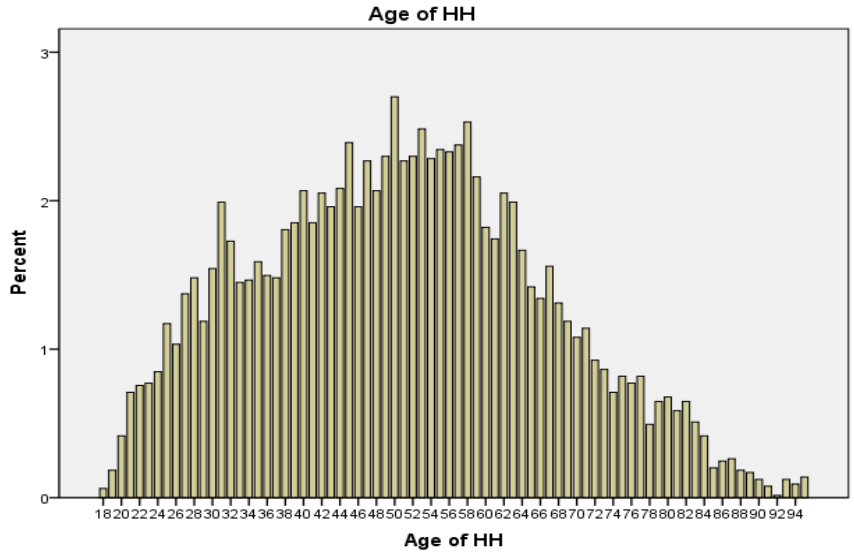
reflect the central tendency of groups. For example, the very large variance in variables such as income and wealth within the same categories highlight the great differences (i.e. inequality) in measures such as income and wealth. Median values are used to characterize the profile of different groups in the overall discussion. However, the mean values of all variables are also given in tables to illustrate that while the mean and median values are similar in some measures such as education, the differences in income and wealth between those groups are sometimes quite large due to variance in those economic measures.

EMPIRICAL RESULTS

Descriptive Statistics

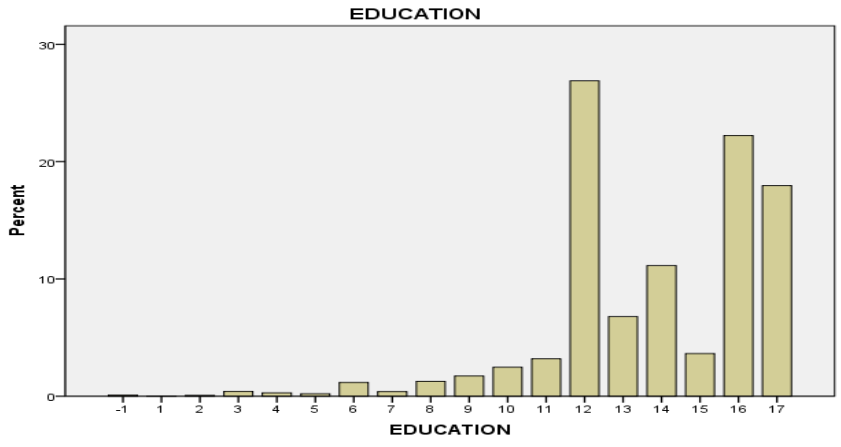
Distributions of age, education, income and net worth will be illustrated in Figures 3, 4, 5 and 6.

Figure 3. Distribution of Age (Years)



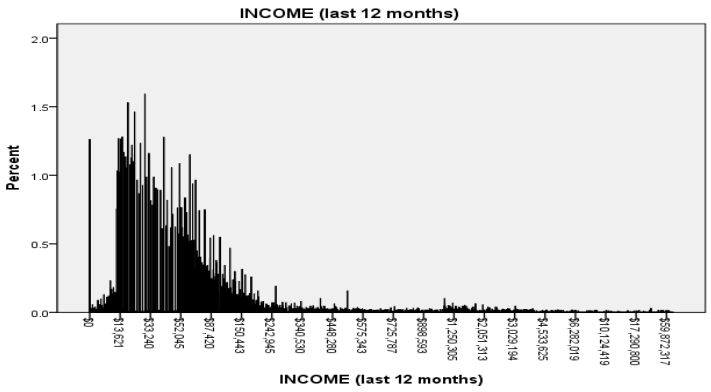
Mean		50.69
Median		50
Mode		50
Std. Dev.		15.88
Skewness		0.197
Kurtosis		-0.54
Range		77
Minimum		18
Maximum		95
Percentiles	20	36
	40	46
	60	55
	80	64

Figure 4. Distribution of Education (Years)



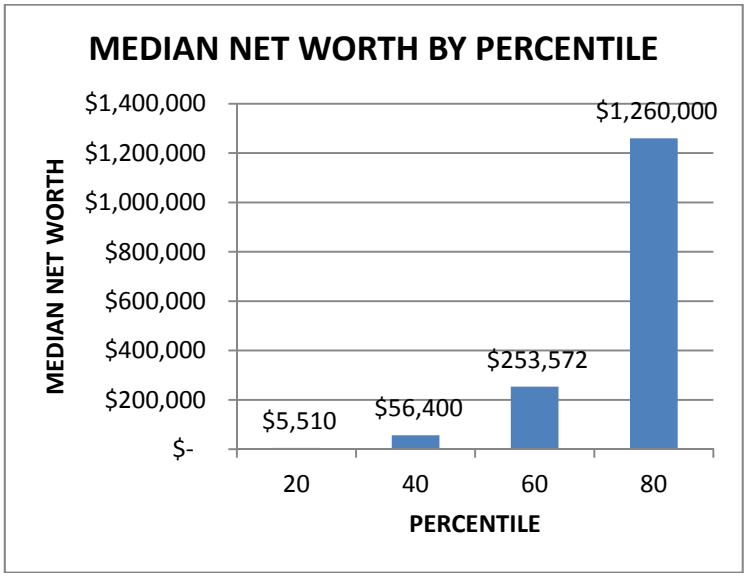
Mean		13.81
Median		14
Mode		12
Std. Dev.		2.739
Skewness		-1.03
Kurtosis		1.881
Minimum		-1
Maximum		17
Percentiles	20	12
	40	13
	60	16
	80	16

Figure 5. Distribution of Income ($)



Mean	$ 612,774	
Median	$ 55,908	
Mode	$ 30,495	
Std. Dev.	$ 6,075,000	
Skewness	$ 35	
Kurtosis	$ 1,646	
Range	$361,368,571	
Minimum	$ -	
Maximum	$361,368,571	
Percentiles 20	$ 22,363	
Percentiles 40	$ 41,677	
Percentiles 60	$ 71,156	
Percentiles 80	$ 158,575	

Figure 6. Distribution of Net Worth ($)



NETWORTH

Mean	$ 7,340,000	
Median	$ 124,355	
Mode	$ -	
Std. Deviation	$ 43,150,000	
Skewness	$ 12	
Kurtosis	$ 183	
Range	$1,000,000,000	
Minimum	$ (6,932,400)	
Maximum	$1,000,000,000	
Percentiles 20	$5,510	
Percentiles 40	$56,400	
Percentiles 60	$253,572	
Percentiles 80	$1,260,000	

Differences Based On *Marital Status* of HH in US Households

Table 1. Household Socioeconomic Profiles based on Marital Status of HH
(Comparison of Married versus Single HH)

MARITAL STATUS		Age (years)	No of Children)	EDUCATION (years)	INCOME (last 12 months)	NET WORTH
MARRIED=Married	Mean	50.96	1.10	14.03	$826,238	$10,500,000
Or LWP (62.8%)	Median	50.89	.83	14.56	$81,227	$268,700
SINGLE=Unmarried	Mean	50.24	.52	13.44	$252,099	$2,000,000
Nor LWP (37.2%)	Median	49.71	.34	13.28	$29,454	$30,020
Total	Mean	50.69	.89	13.81	$612,774	$7,340,000
(100%)	Median	50.51	.61	13.93	$55,754	$124,355

Summary of findings based on Differences in HH Marital Status

Based on the results reported in Table 1, the following comparisons highlight the differences based on *marital status* of HH. Table 2 summarizes the test for statistical significance for differences in median and mean values.

1. Age: *single HH* are younger (median age=49.7 vs. 50.9 years) than *married HH* by 2.2 years.
2. Education: *single HH* are less educated (median education=13.3 vs. 14.6 years) than *married HH* by 1.3 years.
3. Number of children: *single HH* have fewer children (mean=0.52 vs. 1.10) *than married HH,* or about 50% less children or a 1:2 ratio.
4. Income: *single HH* have less income (median=$29,400 vs. $ 81,200) than *married HH*, So, *single HH* median income is approximately 36% of *married HH* median income
5. Net worth or wealth: *single HH* have less *net worth* or wealth (median=$30,000 vs. $268,700) for *married HH*, So, *single HH wealth* is approximately 11% of *married HH wealth* or net worth.

Table 2 summarizes the results of non-parametric testing of differences between median values based on *marital status,* and also the differences in mean values (t-tests). For hypothesized differences, the results reveal statistical significance for 1) age of HH, 2) education, 3) number of children, 4) income, 5) wage income, and 6) net worth.

Table 2. Results of non-parametric test of median differences and mean differences (t-test) based on *Marital Status*				
VARIABLE	MEDIAN COMPARISON RESULTS	SIG.	MEAN COMPARISON RESULTS	SIG.
1. AGE	Single < Married	Yes	Single < Married	Yes
2. EDUCATION	Single < Married	Yes	Single < Married	Yes
3. NUMBER OF CHILDREN	Single < Married	Yes	Single < Married	Yes
4. INCOME	Single < Married	Yes	Single < Married	Yes
5. WAGE INCOME	Single < Married	Yes	Single < Married	Yes
6. NET WORTH	Single < Married	Yes	Single < Married	Yes

Differences in Socioeconomic Profiles of US Households based on Gender of *HH*

Table 3.Comparison of Household Profiles based on *Gender* (Male vs. Female HH)							
Gender of HH (Head of Household)			HH Age (years)	No. Children (inclusive)	Years HH EDUCATION	INCOME (last 12 months)	NET WORTH
	Male 77%	Mean	50.4	.93	14.0	$ 779,759	$ 9,373,940
		Median	**50.4**	**.65**	**14.4**	**$ 69,534**	**$ 196,280**
	Female 23%	Mean	51.5	.73	13.3	$ 55,263	$ 558,410
		Median	**50.8**	**.50**	**13.1**	**$ 26,853**	**$ 23,250**
	Total 100%	Mean	50.7	.89	13.8	$ 612,774	$ 7,342,098
		Median	50.5	.61	13.9	$ 55,755	$ 124,355

Summary of Findings based on Differences in HH Gender

Based on the results reported in Table 3, the following comparisons highlight the differences based on *Gender of HH*. Table 4 summarizes the test for statistical significance for differences in median values.

1. Age: *female HH* are slightly older (median age=50.8 vs. 50.4 years) than *male HH* and the difference is statistically insignificant (Table 3).
2. Education: *female HH* are less educated (median education=13.1 vs. 14.4 years) than *male HH* by 1.3 years.
3. Number of children: *female HH* have fewer children (median=0.5 vs. 0.65) compared to *male HH*.
4. Income: *female HH* have less income (median=$26,900 vs. $ 69,500) than *male HH*, So, female *HH* median income is approximately 39% of *male HH* median income.
5. Net worth or wealth: *female HH* have less *net worth* or wealth (median=$23,300 vs. $196,300) for *male HH*, So, *female HH wealth* less than 12% of *male HH wealth* or net worth.

Table 4. Results of non-parametric test of median differences and mean differences (t-test) based on *Gender*				
VARIABLE	MEDIAN COMPARISON RESULTS	**SIG.**	MEAN COMPARISON RESULTS	**SIG.**
1. AGE	No sig. difference	**No**	Female < Male	**Yes**
2. EDUCATION	Female < Male	**Yes**	Female < Male	**Yes**
3. NUMBER OF CHILDREN	Female < Male	**Yes**	Female < Male	**Yes**
4. INCOME	Female < Male	**Yes**	Female < Male	**Yes**
5. WAGE INCOME	Female < Male	**Yes**	Female < Male	**Yes**
6. NET WORTH	Female < Male	**Yes**	Female < Male	**Yes**

THE COMBINED INFLUENCE OF MARITAL STATUS AND GENDER

Table 5 illustrates the differences when *gender of HH* is introduced as second classification variable. Based on Figure 2 less than 1% of *married HH* has a *female HH*. In some instances, the magnitude of differences is quite stark.

Table 5. The Simultaneous Influence of *Marital Status* and *Gender of HH*						
HH GROUP% OF N		AGE of HH	# of Children	EDUCATION	INCOME	NET WORTH
MARRIED Male (62.2 %)	Mean	51.0	1.11	14.0	$ 832,954	$ 10,595,386
	Median	51.0	.83	14.6	$ 81,441	$ 271,244
MARRIED Female (0.6%)	Mean	44.2	.73	13.6	$ 93,867	$ 826,731
	Median	44.0	.52	14.2	$ 35,324	$ 15,570
MARRIED TOTAL (62.8%)	Mean	51.0	1.10	14.0	$ 826,238	$ 10,506,623
	Median	50.9	.83	14.6	$ 81,228	$ 268,700
SINGLE Male (14.7%)	Mean	48.0	.20	13.7	$ 554,532	$ 4,202,340
	Median	48.1	.14	13.7	$ 34,476	$ 42,600
SINGLE Female (22.5%)	Mean	51.7	.73	13.3	$ 54,283	$ 551,596
	Median	51.1	.50	13.0	$ 26,739	$ 23,475
SINGLE TOTAL (37.2%)	Mean	50.2	.52	13.4	$ 252,099	$ 1,995,231
	Median	49.7	.34	13.3	$ 29,454	$ 30,020
Male TOTAL 76.9%	Mean	50.4	.93	14.0	$ 779,759	$ 9,373,940
	Median	50.4	.65	14.4	$ 69,534	$ 196,280
Female TOTAL 23.1%	Mean	51.5	.73	13.3	$ 55,263	$ 558,410
	Median	50.8	.50	13.1	$ 26,853	$ 23,250
Total SAMPLE	Mean	50.7	.89	13.8	$ 612,774	$ 7,342,098
100%	Median	50.5	.61	13.9	$ 55,755	$ 124,355

In *married HH* households the following comparisons are observed based on *gender of HH*:

1. Age: *married-female HH* are younger (median age= 44.0 vs. 51.0 years) than *married-male HH* by 7 years, which is about 14% younger.
2. Education: *married-female HH* have less education (median education=14.2 vs. 14.6 years) compared to *male-married HH* by 0.4 years or about a 3% difference.
3. Number of children: *married-female HH* have fewer children (mean=0.73 vs. 1.10) than *married-male HH*, so an approximately 2:3 ratio of children, which is about 33% less.

4. Income: married-female *HH* have less income (median=$35,300 vs. $81,400) than *married-male HH*. So, *married-female HH* median income is approximately 43% of *married male-HH* median income.
5. Net worth of wealth: *married-female HH* have less *net worth* or wealth (median=$15,600 vs. $271,200) than *married-male HH*, So*, married-female HH* median *net worth* or wealth is less than 6.0% of married-male HH median net worth or wealth.

In *single HH* households the following comparisons are observed based on *gender of HH*:

1. Age: *single-female HH* are older (median age= 51.1 vs. 48.1 years) than *single-male HH*, or 6% older.
2. Education: *single-female HH* have less education (median education=13.1 vs. 13.7 years) compared to *single-male HH* by 0.6 years. This is about 4% less education.
3. Number of children: *single-female HH* have more children (mean=0.73 vs. 0.20) than *single-male HH*. So, *single-female HH* have approximately 3.7:1 ratio of children, or almost 4 times the number of children for *single-male HH*.
4. Income: *single-female HH* have less income (median=$26.700 vs. $34,500) than *single-male HH*. *Single-female HH* median income is approximately 77% of single-male HH median income.
5. Net worth or wealth: *single-female HH* have less *net worth* or wealth (median=$23,500 vs. $42,600) than *single-male HH*. So, *single-female HH* median net worth or wealth is approximately 55% of *single-male HH* median *net worth* or wealth.

Figure 7. Comparative Summary of Socioeconomic Profiles of Four HH Groups Based on Marital Status and Gender

GROUP 1=MM **MARRIED Male** *(62.2% of sample)*	**GROUP 2=MF** **MARRIED Female** *(0.6% 0f sample)*	**GROUP 3=SM** **SINGLE Male** *(14.7% of sample)*	**GROUP 4=SF** **SINGLE Female** *(22.5% of sample)*
Age = 51.0 years Children= 1,11 (mean) Education= 14.6 years Income= $ 81,400 Net Worth= $ 271,200	Age = 44.0 years Children= 0.73 (mean) Education= 14.2 years Income= $ 35,300 Net Worth= $ 15,600	Age = 48.1 years Children= 0.20 (mean) Education= 13.7 years Income= $ 34,500 Net Worth= $ 42,600	Age = 51.1 years Children= 0.73 (mean) Education= 13.0 years Income= $ 26,700 Net Worth= $ 23,500
Highest income Highest wealth Most educated Highest number of children 2nd oldest group ***** Most Favorable Economic profile of 4 groups.**	Much less income than MM; close to SM Much less wealth than MM Less wealth than even SM 2nd highest Education Children 2nd and tied with SF	SM Single Male HH is better off than Single Female and close to or better off than Married Female HH. Least children	Worst economic profile Single has lowest income and wealth. Also is the oldest group 2nd highest group with child; tied with MF Lowest education ***Worst Economic Profile of 4 groups

SUMMARY OF OLS REGRESSION MODELS

Table 6 summarizes the OLS regression models for *wage income* (Model 1) and *income* HH (Model 2).

Table 6. OLS Regression Model Results for Wage Income and Income

Y= WAGE INCOME (1) Y=INCOME (2)	WAGE INCOME (MODEL 1)		INCOME (MODEL 2)	
X_i	t	Sig	t	sig
X_1 EDUCATION	10.8	0.000	13.9	0.000
X_2 AGE	3.3	0.001	7.3	0.000
X_3 GENDER	-2.5	0.011	-4.4	0.000
X_4 MARITAL STATUS	-2.5	0.012	-1.8	0.071
Adjusted R^2	**0.006**		**0.01**	
F statistic	**47.1**	**0.000**	**85.6**	**0.000**

Influences on Wage Income (Model 1)

$$Y_{\text{(Wage Income)}} = a + bX_{1\text{ (Education)}} + cX_{2\text{(Age)}} + dX_{3\text{(Gender)}} + eX_{4\text{(Marital Status)}}$$

The OLS regression model for *wage income* (Model 1; Table 6) evaluates the influence of education, age, gender and marital status of HH. All four demographic variables have a significant influence or correlation with *wage income*. The relative influence is as follows: education has the strongest positive (t=10.8), and age is 2nd (t=3.3); both gender (t= -2.5) and marital status (t= -2.5) have significant, but negative influences on *wage income*. So, more educated and older HH are correlated to higher *wage income*, but female and single HH are correlated to lower *wage income*.

Influences on HH Income (Model 2)

$$Y_{\text{(Income)}} = a + bX_{1\text{ (Education)}} + cX_{2\text{(Age)}} + dX_{3\text{(Gender)}} + eX_{4\text{(Marital Status)}}$$

The OLS regression model for HH *income* (Model 2; Table 6) evaluates the influence of education, age, gender and marital status of HH. Three out of four demographic variables have a significant influence or correlation with HH *income*. The relative influence is as follows: education has the strongest positive (t=13.9), and age is 2nd (t=7.3); gender has a negative and significant influence on HH *income* (t= -4.4). Marital status also has a negative correlation with *income* but it is not significant (t= -1.8). So, more educated and older HH are correlated to higher HH *income*, but female HH is correlated to lower *income*. The insignificant influence of marital status may be due to the high correlation between gender and marital status. Less than 1% of married households have a female HH, and more than 60% of single households have a female HH (Figure 2).

Table 7 summarizes the OLS regression models for *Net Worth* (Model 3, 4, & 5) using different configurations of explanatory variables.

Influences on Net Worth (Model 3)

$$Y_{(\text{Net Worth})} = a + bX_{1\,(\text{Education})} + cX_{2(\text{Age})} + dX_{3(\text{Gender})} + eX_{4(\text{Marital Status})}$$

Table 7. OLS Regression Model Results for Net Worth

Y (NET WORTH) = f (X_i + …..	**MODEL 3 Net Worth = f (Demographics)**	**MODEL 4 Net Worth = f (Demographics + Wage Income)**	**MODEL 5 Net Worth = f (Demographics + Income)**
X_i	t	t	T
X_1 EDUCATION	22.7 ***	20.6 ***	18.1 ***
X_2 AGE	21.2 ***	21.1 ***	20.4 ***
X_3 GENDER	**-5.4 ****	**-4.9 ****	**-3.6 ****
X_4 MARITAL STATUS	**-7 ****	**-6.6 ****	**-7.1 ****
X_5 INCOME	***omit***	***omit***	109.4 ***
X_6 WAGE INC	***omit***	48.2 ***	***omit***
Adjusted R^2	**0.04**	**0.10**	**0.30**
F statistic	**333**	**625**	**2,758**

Model 3 evaluates the influence of education, age, gender and marital status of HH. All four demographic variables have a significant influence or correlation with *net worth*. The relative influence of the demographic variables on *net worth* is as follows: both education and age have a positive influence, but education (t=22.7) is more significant that age (t=21.2). Marital status and gender are both significant and negative, but marital status (t= -7.0) is more significant than gender (t= -5.4). The adjusted $R^{2\,\text{for}}$ Model 3 is 0.04 (4 % explanation of model and F-statistic of 333), so the demographic variables account for about 4% explanation of *net worth* is still significant, although the amount of explanation is small (4%).

Influences on Net Worth (Model 4)

$$Y_{(\text{Net Worth})} = a + bX_{1\,(\text{Education})} + cX_{2(\text{Age})} + dX_{3(\text{Gender})} + eX_{4(\text{Marital Status})} + f\mathbf{X_{5(\text{WAGE INCOME})}}$$

Model 4 evaluates the influence of education, age, gender and marital status of HH in addition to *wage income*. *Wage income* has the strongest positive influence on *net worth* (t= 48.2). All four demographic variables have a significant influence or correlation with *net worth*. The relative influence of the demographic variables is as follows: both education and age have a positive influence, but age (t=21.1) is more significant that education (t=20.6). Marital status and

gender are both significant and negative, but marital status (t= -6.6) is more significant than gender (t= -4.9).

The adjusted R^2 for Model 4 is 0.10 (10 % explanation of model and F-statistic of 625), so the model's explanation of *net worth* is more significant than Model 3. By introducing *wage income* to Model 4, the explanation is increased by 6% compared to Model 3.

Influences on Net Worth (Model 5)

$$Y_{(\text{Net Worth})} = a + bX_{1\,(\text{Education})} + cX_{2(\text{Age})} + dX_{3(\text{Gender})} + eX_{4(\text{Marital Status})} + i\mathbf{X_{6(INCOME)}}$$

Model 5 evaluates the influence of education, age, gender and marital status of HH in addition to Income (during last 12 months, which includes Wage Income). HH income has the strongest positive influence on *net worth* (t= 109.4). Wage income is deleted from this model, so it doesn't confound the effect of HH income. All four demographic variables have a significant influence or correlation with *net worth*. The relative influence of the demographic variables is as follows: both education and age have a positive influence, but age (t=20.4) is more significant that education (t=18.1). Marital status and gender are both significant and negative, but marital status (t= -7.1) is more significant than gender (t= -3.6).

The adjusted $R^{2\,\text{for}}$ Model 4 is 0.30 (30 % explanation of model and F-statistic of 2,758), so the model's explanation of *net worth* is very significant. So, the introduction of Income increases by 26% compared to Model 3 and 20% compared to Model 4.

SUMMARY CONCLUSIONS AND IMPLICATIONS

Based on the regression models, it is obvious that while demographics such as education, age, gender and marital status have significant correlations with wage income, HH income and net worth, they may not be the most significant explanatory variables for understanding income and wealth inequalities. Also, the correlation between explanatory (independent) variables can distort the regression results.

However, the objective of the regression models was to give an overview of the relative importance or influence of explanatory variables.

Model 1-Wage Income: Education and age have positive correlations with wage income; so more educated and older HH have higher wage income. Gender and marital status have negative correlations with wage income (female and single HH are correlated to lower wage income). So, the relative order of correlation is as follows:

WAGE INCOME: (+) Education > (+) Age > (-) Gender, (-) Marital Status

Model 2 –Income : Education and age have positive correlations with Income; so more educated and older HH have higher income. Gender and marital status have negative correlations with income, but only gender is significant (female HH is correlated to lower income). So, the relative order of correlation is as follows:

INCOME: (+) Education > (+) Age > (-) Gender (-) Marital Status but n.s.

Model 3-Net Worth: Education and age have positive correlations with net worth; so more educated and older HH have higher net worth. Gender and marital status have negative correlations with net worth (female and single HH are correlated to lower net worth). Marital status has a greater negative effect than gender with respect to net worth. So, the relative order of correlation is as follows:

NET WORTH: (+) Education > (+) Age > (-) Marital Status > (-) Gender

Model 4-Net Worth: Wage income is included in Model 3. Wage income is the dominant explanatory variable, compared to similar demographics (Model 3).

NET WORTH: (+) Wage income > (+) Age > (+) Education > (-) Marital Status> (-) Gender

Model 5-Net Worth: Income is included in Model 3. Income is the dominant explanatory variable, compared to similar demographics (Model 3), and also Model 4.

NET WORTH: (+) Income > (+) Age >(+) Education > (-) Marital Status > (-) Gender

While the regression models are helpful in understanding relative effects of a number of variables simultaneously on wealth and income, they may also be confounded by multicollinearity effects in the models. The combination of models used was an attempt to separate out and compared relative overall effects. The conclusion is that gender and marital status are correlated to wealth and income variables, but they may be more important as proxies for other societal and socioeconomic structural issues.

The results of the comparison between the four groups (based on gender and marital status) provide a simple, but important starting point for understanding the differences in socioeconomic groups. Overall, *married HH* families have significantly higher incomes and wealth than *single HH* families. However, when gender of HH is introduced the results change dramatically. For example, *married male HH* (Group 1) still have significantly higher incomes (about $81,000) and wealth (about $270,000), but there are significantly lower median values for *married female HH (Group 2)* income (about $35,000) and wealth (only about $16,000). Furthermore, *married female HH* (Group 2) income is closer to *single male HH* (Group 3) ($35,000) and has less wealth than Group 3's $42,000 median value. The lowest income profile is *single female HH* (Group 4) (income= about $28,000) and median wealth is around $24,000. The lowest wealth group overall is the *married female HH* (median wealth=$16,000).

Both groups headed by women have the 2nd highest number children (0.73) compared to the highest *married male HH* (1.11) and lowest *single male HH* (0.20). What are the implications for the *female HH* groups where income are lower, but the presence of children is significant? It may be particularly distressing *single female HH*, which constitute 22.5% of all households, show a mean of 0.73 children and a median income of less than $28,000. What are the implications for lack of resources for educating and raising children and circumventing inter-generational propagation of poverty?

Continued research can add more depth to our understanding of the structural and fundamental issues which may be more significant. It is also important to explore how the

combined interaction of variables influences differences. After all, households do not live in a vacuum, where single variables have well-defined and uncontaminated effects.

REFERENCES

Diaz-Giménez, J., Glover, A., & Ríos-Rull, J. (2011). Facts on the distributions of earnings, income, and wealth in the United States: 2007 Update. *Quarterly Review*, 34(1), 2-31.

Esping-Andersion, G. (2007) Sociological explanations of changing income distributions. *American Behavioral Scientist, 50*, 639-58.

Fry, R. & Cohn, D. (2010). Women, men and the new economics of marriage. Pew Research Center. Accessed September 10, 2013. URL http://www.pewsocialtrends.org/files/2010/11/new-economics-of-marriage.pdf

Gustman, A. L., & Juster, F. T. (1995). *Income and wealth of older American households: modeling issues for public policy analysis* (No. w4996). National Bureau of Economic Research.

Holden, K. C., & Kuo, H. H. D. (1996). Complex marital histories and economic well-being: The continuing legacy of divorce and widowhood as the HRS cohort approaches retirement. *The Gerontologist*, *36*(3), 383-390.

Lusardi, A., & Mitchell, O. S. (2007). Baby boomer retirement security: The roles of planning, financial literacy, and housing wealth. *Journal of monetary Economics*, *54*(1), 205-224.

McCall, L., & Percheski, C. (2010). Income inequality: New trends and research directions. *Annual Review of Sociology*, *36*, 329-347.

McLanahan, S. (2004). Diverging destinies: How children are faring under the second demographic transition. *Demography*, *41*(4), 607-27.

Neill Hoch I., & Mohan-Neill, S. (2013). Can financial literacy help alleviate wealth and income inequality in the US? *Allied Acadmies: Academy of Marketing Studies Proceedings, 18*(1) 31-36.

Norton, M. I., & Ariely, D. (2011). Building a better America—one wealth quintile at a time. *Perspectives on Psychological Science, 6*(1), 9-12.

Seigel, J. S. (1993). *A generation of change: A profile of America's older population*. New York: Russell Sage Foundation

Wilmoth, J., & Koso, G. (2002). Does marital history matter? Marital status and wealth outcomes among preretirement adults. *Journal of Marriage and Family*, *64*(1), 254-268.

Wolff, E. N. (1992). Changing inequality of wealth. *The American Economic Review*. 552-558.

Wolff, E. N. (1998). Recent trends in the size distribution of household wealth. *The Journal of Economic Perspectives*, *12*(3), 131-150.

RECONSTRUCTING LONG-RUN ECONOMICS: SURVEY AND ISSUES

Oluremi Ogun,University of Ibadan,

ABSTRACT

This article surveys the literature on the usages and applications of the notion of long-run in economics. Issues therefrom, are discussed and probable solutions proffered to thorny issues. The developing professional economist is advised to always resist the temptation to model every economic phenomenon in terms of the long-run. In some branches of economics, the focus of policy is often the short run and this should be respected in the modeling process. Where it is deemed absolutely necessary to undertake a long-run exercise, attention should be paid to issues of model specification and analysis.

Key Words: General Economics; Economic Education and Teaching of Economics; History of Economic Thought; Economic Methodology, Econometric Modeling; General Aggregative Models; Technological Change, Research and Development; Economic Growth and Aggregate Productivity.

JEL Nos.: A1; A2; B1; B21; B22; B23; B26; B41; C50, C51; C58; O3; O4

INTRODUCTION

In the literature, the origin of long-run in economic thinking and analysis appears to date to the Classical economists, in particular, David Hume, in his checkmating drive against the Mercantilists around the mid eighteenth century. There is perhaps no area of economics that is today immune to the long-run influence. Over time however, several economists of the classical brand and those of other persuasions have conceptualized the notion variously in the literature. The beginning student is told in microeconomics that the long-run curve could in some contexts equal the short-run curve or could be an average of the minimum points on all the short run curves. He soon discovers that there are somewhat related but computationally different counterparts in macroeconomics – permanent income, and in financial/monetary economics – term structure of interest rate.

At a later date in the course of his training, he is told that some real variables (deflated or natural) e.g. saving and(fixed) capital formations generate long-run effect. He begins to believe that all variables have long-run or steady state values. Just then, he is introduced to the concept of natural rate level of output and the associated long-run aggregate supply curve and told that only factors/variables causing shift of this curve qualify to be admitted into the class of long-run growth drivers. Not only does he begin to view variables such as, technology (National System of

Innovations), tastes and productivity growth as the critical drivers in this sense, but also has written in his memory, an economy-wide or macroeconomy long-run state.

His confused state knows no bound when as a research student, he traverses the literature on economic growth, demand for money, balance of payments and productivity growth and encounters varying application of the notion of long-run. He clearly needs help to separate the grain from the chaff in the course of his professional growth; this is the essence of this survey. It is generally non-technical and does not pretend to be exhaustive of the subject but at least, endeavors to pinpoint key issues of emphasis, focus and the attendant conceptual and application problems in the literature on the different models.

The rest of the paper is organized as follows. Section II deals with growth models as evolved dynamically while section III is concerned with financial models emphasizing long-run effect. Section IV focuses on econometric approaches to long-run modeling while section V deals with the matters arising from the review, drawing support from the writer's research endeavors and those of some of his graduate students. Section VI provides some concluding remarks.

Economic Growth

Classical economists such as Smith (1991), Malthus (1798) and Ricardo and Eck (1817) did not directly address the issue of long-run growth though occasional references to 'steady state' permeate their expositions on growth. Mostly, particularly, Smith, Malthus and Ricardo emphasized the efficacy of the 'invisible hand' in restoring the economy of the typical capitalist economy to long-run equilibrium. All of them recognize the importance of productivity (of labor – which depends on capital investment – in the case of Smith and Malthus, and of capital, in the case of Ricardo) in the drive up to the long-run. While Smith insinuated about the possibility of the economy growing continuously and powered by increasing capital investment driving labor productivity, both Malthus and Ricardo were pessimistic outright about the possibility of the economy growing beyond the 'steady state (Both Malthus and Ricardo believed that labor productivity would be counterbalanced by population growth in the steady or stationary state). Even, the mathematical formulation of the Classics' steady state condition for increasing capital accumulation by Ramsey (1928) and later applied by other neoclassical economists such as, Cass (1965) and Koopmans (1965)does not appear to have improved on the submission of the Classics; the odds against growing capital accumulation beyond the steady-state are simply very weighty (They submit that, the steady-state capital stock will be higher if capital is more productive and lower if consumers are more impatient, population growth is faster, depreciation is greater, or technological progress is more rapid).

The original classical growth models were succeeded by the Keynesian models as pioneered by Harrod (1939) and Domar (1946). The Keynesian models mostly used production functions with little substitutability among the inputs to argue that the capitalist system is inherently unstable. Since these arguments were developed during or immediately after the Great Depression, they were received sympathetically by economists. However, they are not relevant to the analysis of long-run as the Keynesians of the time did not believe in the idea (More recent Keynesian models generally admit the existence of long-run in the economy).

The next and most important contribution to modern growth theory have been the works of Solow (1956) and Swan (1956) both of neoclassical orientation. The distinctive feature of their

model which is also referred to as the exogenous growth model is in its special specification which is a production function built on the assumptions of constant returns to scale, constant saving rate, diminishing returns to factor inputs and some positive and smooth elasticity of substitution between the inputs. The assumption of factor substitutability is to allow for a stable equilibrium growth and the constant saving rate makes possible, the generation of a simple general equilibrium model of the economy.

A key prediction of the aforementioned neoclassical model is conditional convergence which interprets simply that, the lower the starting level of per capita gross domestic product (GDP) compared with the long-run or steady-state position, the faster the growth rate. Such a possibility is induced by the assumption of diminishing returns to capital according to which economies with low capital per worker tend to generate relatively higher rates of return and higher growth rates. The convergence is conditional in the Solow-Swan model because steady-state levels of capital and output per worker depend, on the saving rate, the growth rate of the population and the position of the production function that might vary across economies. Recent empirical studies have suggested additional sources of cross-country variation, particularly, differences in government policies and in initial stocks of human capital.

In terms of the long-run, the Solow-Swan model basically implies that a higher saving-investment ratio simply raises the level of income per capita but would have no effect on the rate of economic growth. This way, per capita growth must eventually come to an end. The only way to alter this equation and avoid such an implied stagnation is to augment the production processes of the economy with new technologies. In other words, unlike Malthus and Ricardo, the model explicitly recognizes the role of technological progress at generating growth beyond the steady-state. Thus, technological innovation received first mention as long-run growth driver. In recognition of this novel submission, the neoclassical economists of the late 1950s and 1960s explicitly factored technological progress as an exogenous variable into their growth models while also retaining the equally novel prediction of conditional convergence (The idea of convergence is actually traceable to Malthus in his analysis of population growth dynamics. However, it seems that the negative implication of his theory for mankind informed the very scanty mention of his convergence proposition in the literature even when some empirical regularity could be adduced to it).

The first known attempt to explicitly introduce ideas into growth models was by Arrow (1962) and later by Sheshinski (1967). In these models, ideas were unintended by-products of production or investment. This mechanism is described as learning-by-doing. The setting is such that, each new idea immediately spreads through the entire economy. This diffusion process might be technically feasible because knowledge is non-rival. This basic neoclassical model was further developed by Cass (1965),Koopsman (1965) and others. In particular, Cass and Koopsman applied Ramsey's analysis of consumer optimization in the drive to endogenize the determination of the saving rate. This extension tends to preserve the hypothesis of conditional convergence while allowing for strong transitional dynamics. However, as discovered by the authors, it is not easy to create compatibility between the theory of technological change and the neoclassical framework because the standard assumptions of competition cannot be met. The reason being that technological progress requires the creation of new ideas which as shown in the work of Arrow

(1962) and Sheshinski (1967), are partially non-rival. Cass (1972) thus represents a refinement of the initial efforts in this area.

Beginning from the early 1970s and for about 15 years, there appears to be an interregnum in long-run growth theorizing in the literature with the advent of rational expectations and on the eve of the oil shocks. Accordingly, short run vacillations dominated most of economic thinking to the extent that the incorporation of rational expectations into business cycle models received the most attention.

Resurgence occurred around the mid-1980s with the entry of endogenous growth models heralded by the pioneering works of Romer (1986) and Lucas (1988).The aim was to study the determinants of long-run growth and put more emphasis on these determinants rather than the mechanics of business cycles or the countercyclical effects of monetary and fiscal policies. However, the recognition of the importance of long-run growth was only a first step. The second step was to abandon the main idea of the neoclassical growth model according to which the long-run per capita growth rate is linked to the rate of an exogenous technological progress (The endogenous growth models are generally referred to as the AK models due to the usual presence of AK in the production function; A, represents the level of technology and K, often, is a composite of human and physical capital. Some of the models e.g. Lucas (1988) however, enter human capital distinctly and independent of physical capital).

The initial phase of the new theory which was based on the research by Arrow (1962), Sheshinski (1967) and Uzawa (1965) as re-modeled by Romer (1986), Lucas (1988) and Rebelo (1991) did not really present a theory of technological change. In these models, growth may continue limitless because, the returns to investment in capital resources, including human capital, do not diminish necessarily as economies grow; human capital and the flow of knowledge across producers help to prevent the tendency for diminishing returns to the capital accumulation. This wave of research which incorporated research and development (R & D) theories and imperfect competition into growth models began with Romer (1987, 1990) and found significant contributions by Aghion and Howitt (1992) and Grossman and Helpman (1991).

In these models, purposive R & D activity which results from some form of ex post monopoly power, results into technological progress. Thus, as long as the economy is not devoid of ideas (new ideas), the growth rate can be raised in the long-run. However, the growth rate and the inventive activity may not be Pareto optimal, due to the failure to create the new product and innovate on the production methods. In this framework, the long-term growth depends on government intervention which could take the form of taxation, maintenance of law and order, provision of infrastructural services, protection of intellectual property rights and regulations of international trade, financial markets and other aspects of the economy. Thus, the long-term growth can greatly be influenced by government actions.

The research agenda of endogenous growth also covers models of diffusion of technology. The diffusion models are related to the way in which follower economies contribute to these advances by imitation of leading-edge economies, while the analysis of innovation deals with the rate of technological progress in these advanced countries. Since imitation is cheaper than innovation, the diffusion models predict a form of conditional convergence that resembles the prediction of the neoclassical growth model. Also, it implicitly underwrites human capital as a

long-run growth driver as countries with an accumulation of such capital tend to absorb new technologies faster (See Barro (1991) citing Nelson and Phelps, 1966).

Another key exogeneity assumption of the neoclassical growth model jettisoned by endogenous growth is the growth rate of the population. To the neoclassic, the higher the growth rate of the population, the lower the steady-state level of capital and the output per capita growth rate for a given initial level of per capita output. This belief does not however, take account of the effects of per capita income and wage rates on population as well as the resources used up in the process of child rearing. Under the endogenous models, population growth is endogenized through the incorporation of an analysis of fertility choice into the neoclassical model. The results obtained from these studies are consistent with the empirical regularity of fertility rates. This is so because, fertility rates tend to fall with per capita income over the main range of experience, but it may rise with per capita income for the poorer countries. Another growth research based on the endogeneity of labor supply is concerned with migration and labor-leisure choice.

In general, the endogenous growth theory has developed into two generations. The first phase was the development of the generation of semi-endogenous growth models and the second was the Schumpeterian growth theory. The most contributors to the semi-endogenous models were Jones (1995), Kortum (1997) and Segerstrom (1998). The key element of this contribution is the abandon of scale effects in ideas generation by assuming diminishing returns to the stock of R & D knowledge. Thus, R & D is assumed to increase continuously to sustain positive total factor productivity (TFP) growth rate.

The second generation was the Schumpeterian models which have been developed by Aghion and Howitt (1994, 1998), Peretto (1998), Young (1998), Dinopoulos and Thompson (1999) and Peretto and Smulders (2002). These models maintain the assumption of constant returns to stock of R & D knowledge. However, they assume that the effectiveness of R & D decline due to the prosperity of products as the economy grows. In general, growth can still be sustainable at a constant level, provided that R and D process is kept to a fixed proportion of the number of product lines. In turn, this is proportional to the size of the population along the balanced growth path. As such, R & D has to rise over time to overcome the increasing range and complexity of products lowering the productivity effects of R & D activity, in order to ensure a sustainable TFP growth rate.

Financial Models

Most generations of classical economists believe that money or finance has no effect on the real sector and thus, it is of negligible long-run value. The debate on the famous classical dichotomy and the associated neutrality/super neutrality of money is very instructive. And, as recent as 1968, Milton Friedman of the Monetarist fame wrote on what monetary policy could do: altering the real equilibrium of the economy was not one of such. By and large, the New Classical Macroeconomic School led by Robert Lucas upholds this tradition of the Classics in its analysis of policy effectiveness under rational expectations.

A somewhat new direction to thinking about the way money or finance affects the macroeconomy began to emerge around the mid-1970s following the simultaneous publications of Mckinnon (1973) and Shaw (1973) on the doctrine of financial repression and its antidote, financial liberalization (Both scholars are traditionally, neoclassical. Note also that, the concept of

financial liberalization dates back a much longer period when it was generally referred to as financial deepening (see e.g. Wallich, 1969). However, the 1973 publications of the Mckinnon and Shaw added fresh insights). Simply put, the doctrine of financial repression identifies the interventionist policies of governments (mostly, in developing countries) as the bane of underdevelopment of the financial sector, manifesting in slow or stunted growth of the real sector of the economy. The culprit policies were identified to include, fixed or administratively determined interest rates in the face of rising/unsteady inflation, fixed exchange rate maintained through frequent reserve intervention in the foreign exchange market, reserve requirements penalizing the banking sector and discriminatory high taxes on financial transactions. While acknowledging that the descent into financial repression may have been inadvertent with financial restriction being the original target, the eventual end result, imposes severe penalty on banks' ability to competitively mobilize saving, take risk and innovate while reducing the volume (in real terms) and quality of credit and hence, investment, in the economy. Correcting these anomalies through a reform of banking and financial sector policies in order to raise the average efficiency of investment would produce a multiplicative impact on growth.

The school of thought that emerged subsequently was to generate several publications demonstrating both theoretically and empirically, the mechanisms by which the new view would affect steady state growth of the typical financially repressed developing economy. Prominent members of this school include Kapur (1976), Galbis (1977), Fry (1978, 1980a, 1980b, 1981), Mathieson (1980), and Yoon Je Cho (1986,1988). In particular, real interest rate, required reserve ratio and bank credit were emphasized as sources of long-run growth (The Neostructuralists particularly, Taylor (1983), Buffie (1984) and vanWijnbergen (1983, 1984) took serious issues with the Mckinnon- Shaw School's submission demonstrating that, the omission to accommodate the informal financial sector of the economy in the various financial models may have been the source of the school's optimistic conclusion on the role of financial sector reforms on growth). In some contexts, a narrative index of financial reform policies is constructed in order to capture both exogenous shift in policy stance as well as the endogenous response of monetary/financial policy to economic development. The resultant series is then plugged into the growth model as a long-run driver. Generally, such narrative index constructions are usually guided by Romer and Romer (1989) and Boschen and Mills (1991).

The Mckinnon-Shaw school's prescription for long-run growth appears to have received some support from the endogenous growth models. Under this extension to the endogenous models (as inspired by Romer, 1990), financial repression is modeled as a disincentive to innovative processes hence, causing productivity growth to fall. Thus, the resultant technological slow down generates adverse effect on steady-state growth. Bencivenga and Smith (1991), De Gregorio (1992), Roubini and Sala-i-Martin (1992) and King and Levine (1993) are some of the major contributors to this literature (In some of these models (e.g. Roubini and Sala-i-Martin, 1992), financial innovation is represented by a distinct variable, that is, in addition to the basic AK features).

Econometric Approaches

Overtime, economists have applied different econometric methods in the investigation of the notion of long-run as may be warranted by the foci of their studies. The observed dynamics in

this regard have been largely informed by developments not only in economics itself but also in econometric theory. For example, the emergence of the unit root – cointegration literature in the past two decades was a response to the perceived defects in the earlier econometric approaches to the issue of long-run. Also, the Lucas (1976) critique on macroeconomic policy formulation procedure resulted in the refinement of existing structural models of the economy and played a key part in the development and popularity of structural vector autoregressive (SVAR) approach to long-run modeling. As noted by Pesaran (1997), long-run estimation in economics is now commonly associated with the notion of equilibrium in the wake of the unit root – cointegration methodology, first introduced by Granger (1981) and elaborated upon for stronger theoretical basis by Engle and Granger (1987). However, Pesaran (ibid.) also observed that, much of the time series long-run analysis is being conducted without explicit account of the underlying equilibrium theories. In what follows, a brief discussion of the major econometric approaches to the analysis of the notion of long-run in the literature is undertaken.

Ordinary Least Squares

The Least squares method (OLS),has been described as the most common approach to long-run modeling (see e.g.Monfort and Mulder, 2000; Mora, 2006; Afonso, Gomes and Rother, 2007)). The method asserts very simply that, to fit a point to the data values, the sum of the squares of the vertical difference from each of the point to the fitted line must be as small as possible in the long-run. The idea of using least squares approach for long-run estimation is to minimize the sum of squared residuals associated with the data, such that, all the endogenous variables will grow at constant (not necessarily identical) rates overtime.

In time series analysis, OLS approach constitutes a veritable means of estimating classical long-run relationships. For instance, the methodology underlies the specification of long-run equilibrium equation for data that are purely non-stationary due to unit roots, but are integrated of the same order. Thus, for example, in the specification below, the error term, μ_t, is believed to be identically and independently distributed (IID).

$$y_t = \beta_0 + \beta_1 x_t + \mu_t \qquad (1)$$

The estimate obtained from the specification is referred to as long-run estimates. By its simple approach, OLS method became endeared to many and has been widely applied in the estimation of long-run phenomena. Kim, Fraser and Hyndman (2007) estimated a long-run response parameter from a general dynamic least squares model, and observed that the interval estimation obtained from the long-run response (elasticity) parameter using a general linear dynamic model was superior to other long-run response estimation methods. The authors observed that this has manifested in three main areas. First, point estimates of the long-run response are often extremely desperate (see also, Marquez and McNeilly, 1988; Askari and Cummings, 1977). Second, it is difficult to estimate accurately the sampling variability of the long-run response estimator (see also, Li and Maddala, 1999). Third, the long-run response estimator can be severely biased when small sample is used. In using interval estimation approach, the small sample biases associated with the parameter estimators are adjusted in the two stages of the bootstrap and it was observed that the outcome performed better for small sample estimates.

In another perspective, Monfort and Mulder (2000), Mora (2006) and Afonso, Gomes and Rother (2007) employed linear method and ordered response models to estimate short and long-run determinants of sovereign debt ratings. The result from the linear method showed a good performance for the estimated model while the ordered probit model differed only in term of the overall predictive power (This has become a natural approach to such type of problem when rating is a discrete variable and reflects the order in terms of the probability of default (see Afonso, *et al*, 2007).). Pesaran and Shin (1996) showed in a similar analysis, but using a general linear autoregressive distributive lag (ARDL) model, that valid asymptotic inferences on the short and long-run parameters can be made using least square estimates. So far, the use of the linear method has been limited by two factors. On the one hand, it is static and does not separate the short-run phenomena from the long-run as in the use of dynamic models such as Structural VAR and Vector Error Correction Models.

On the other hand, the use of only a set of exogenous regression equation that sets some factors to influence the dependent variable within the framework of single-equation has been considered to be faulty (Kibritcioglu and Dibooglu, 2001). Such studies usually consider one-way causalities running from selected economic (and recently, non-economic) regressors to a particular determined variable, ignoring the possible endogeneity of the variable factors. It is also noted in the study by Kibritcioglu and Dibooglu, above that, while long-run modeling is strictly within the confine of the use of OLS methodology, the issue of long-run economic growth may sometime be more complex than to be captured by estimating only single-equation regressions. Attempt at overcoming the problem is the use of log linear specifications to represent long-run situations (Al-Azzam and Hawdon, 1998). One major advantage of this approach is that it yields elasticities in a convenient form and has been found to work well with studies using error correction methods.

Two Stage Least Squares

Two stage least squares (2SLS) modeling approach is used to estimate parameters of identified structural equations. The methodology arose as an attempt to correct simultaneity biases in structural equations models. In the presence of simultaneity problem, the methodology of OLS will fail to produce robust estimates, as the resulting endogenous variable is likely to be correlated with another error term. Simultaneity problem arises when some of the regressors of a model become also endogenous. Based on the underlying assumptions of OLS methodology, the implication of such resultant situation is that there will be a correlation between the endogenous variables and another random error term.

Two stage least squares involves two stages of regressions. The first stage of the regressions attempt to get rid of the resulting correlation between the endogenous variables and the resulting random error terms. The first regression runs all the explanatory variables in the system against the exogenous variable. The second stage involves replacing exogenous parameter with its estimated value, and then applying a regression to the transformed equation. The basic idea of 2SLS is to correct the stochastic error term in the first equation of the influence on the explanatory variables with which it is correlated. The resulting estimates obtained will be consistent as well as normal (especially for large samples). The main advantage of 2SLS approach to econometric modeling is that each of the equations is/can be treated and estimated separately. The 2SLS is just an OLS specification adjusted to correct simultaneity problems. A long-run

specification of 2SLS is usually in semi-log (log-ln). In a situation where the right hand side of the equation becomes endogenous thereby, correlating with the error term, OLS estimations become imprecise. This is the reason for two stage regression approach.

The 2SLS approach requires that the long-run equilibrium relationship (i.e. the cointegrating regression) be modeled as a simple regression involving the levels of the variables. In the first step, all dynamics are ignored and the cointegrating regression is estimated by the OLS. Since the variable in such model is likely to be spurious, only a little attention should be paid to standard error estimates (and thus, t-statistics) in the cointegrating regression. This first regression however serves as a veritable source of comparison with the second regression after being corrected.

Large-Scale Simultaneous Equation Models

The use of Large-Scale Simultaneous Models (LSEM) became popular following a number of models that were developed by the Federal Reserve Board in US and other versions in UK and Australia (Brayton and Mauskopt, 1985; Murphy 1992 and Brayton and Tinsley, 1996). Two factors however account for the little popularity of the models in the recent time. First, there was the relatively poor forecasting of the models in the face of stagflation that manifested in many economies in the 1970s. Second, the advent of rational expectations economics was a major setback for the LSEMs especially, as regards the reliability of their forecasts. For example, the Lucas critique of 1976, and subsequently, Sims (1980b) critique on identification raised serious doubts about the usefulness of the models. The critique generated interests on the use of VAR for macro-econometric analysis. According to Garratt *et al* (2000), this development manifested in three important areas of macro modeling. First, its influence in cointegration analysis arising partly from the response to Sims' critique of the use of incredible identification restrictions involving short-run dynamics which subsequently, led to a consensus that the most important aspect of any structural model is its long-run relationships. Thus, a model must be identified without having to restrict its short-run dynamics. Second, in response to the criticism that large-scale models paid insufficient attention to the micro foundations of the underlying relationships and the properties of the macroeconomic system, economic theory has become better patronized in the specification of large models. Third, as a resolution to Lucas critique, rational expectations (RE) theory becomes incorporated into large-scale models.

Maximum Likelihood Estimation

The maximum likelihood approach (MLE) is used for any function that is conditioned on a set of sample outcomes. It is possible that, in the course of modeling a situation, the dependent variable may be bounded (such as between zero and one), implying that the use of the OLS methodology will be unsuitable as their resulting estimate will be biased. Maximizing the likelihood function (usually with respect to a particular value) implies finding the value of such variable in question. This amounts to maximizing the probability of obtaining the sample values that have actually been observed. This approach has been used by a number of long-run analysts. For example, Patibandla (1996) employed a Cournot Oligopoly model, and used the MLE to solve for the profit maximizing conditions for domestic sales and exports equilibrium in an import restricted market.

Cointegration Approach

Cointegration analysis is a technique applied to estimate the accuracy of correlations between two or more non-stationary variables (Amiri *et al.*, 2011). If the trending time series involved in the regression tends to move together, or are of the same wave-length, this suggest that the use of OLS regression to estimate such model is not likely to be spurious. The existence of cointegration between, say, two macroeconomic variables implies "a true long-run economic relationship" which prevents the residuals (of the resulting equation) from becoming increasingly larger in the long-run.

The use of cointegration analysis has played a key role in economics since its inception. In many instances, it has been referred to as the most revolutionary development in econometrics since the mid-1980s (Seddiighi, 2000). Before then, most economists use to apply linear regressions on non-stationary time series data which are known to produce spurious relationships (Granger, 1981).The essence of testing for cointegration is to avoid spurious regression estimates. The synchrony of the non-stationary time series is the idea behind the concept of cointegration (Gujarati, 2004). Two variables are said to be cointegrated if they have a long-term, or equilibrium relationship between them.

It is important to mention that, cointegration approach may produce different result from other methodologies mentioned earlier. For instance, Sinha (1999) employed a cointegration technique to estimate a traditionally formulated import demand function which ignored the impact of foreign exchange rationing and other restrictions, and did not find any relation between foreign exchange availability and import in India. This result is in contrast with the long-run income and price elasticity estimated by Emran and Shilpi (2008) with a model satisfying the theoretical signs and restrictions. And, the results obtained from the latter study were found to be both significant and economical.

Another major advantage of the cointegration technique is that, it also provides applied econometricians with effective formal framework for testing and estimating long-run models from actual time-series data (Utkulu, 1994). The Engle and Granger static type of long-run ordinary-least-squares regression parameters eventually became accepted as both consistent and highly efficient (Stock (1987)). This position was however not unchallenged (see Banerjee *et al*. (1986) and Blough (1988)). The argument is that ignoring the lagged terms in small samples is likely to create a bias in the estimated parameters. This criticism may have prompted methodology which either tries to incorporate dynamic components (in the form of differencing or lags), or that is concerned with appropriate corrections and modifications to the static parameter estimates. The aftermath of the two responses is that since the two groups of critics emphasize different aspects of the problem, they naturally lead to different solutions. One of the eventual results to the evolving debate was the adoption by many authors of the error-correction estimator approach.

Vector Autoregressive Approach

The Vector Autoregressive (VAR) model, pioneered by Sims (1972, 1980a), is a general framework to describe the dynamic interrelationships between stationary variables. In its original sense, the VAR methodology is used to observe the interdependencies among short-run variables. The modeling approach advocated for long-run relationship is the structural cointegrating VAR

approach. It is based on log-linear model estimated subject to long-run relationship obtained from economic theory. Pesaran (1997) observed that for the purpose of empirical analysis, it is suitable to fit the short-run relation of variables within a suitable multivariate model such as VAR with unrestricted coefficients. By implication, the long-run relation of such model can be embedded within a restricted VAR. Similar approaches of long-run methodology has been suggested using the global VAR (GVAR) model of Pesaran*et al* (2004), and which was further developed by Dees *et al* (2007).

The foregoing implies that the dynamics of adjustment to long-run equilibrium in any of such model must have been restricted. This can be achieved by utilizing the intertemporal nature of the underlying optimization problem as it is done under the rational expectations hypothesis or from the specifications of economic theory. The implication of this modeling approach is that it is relatively simple to implement for optimization problems with linear constraints and quadratic objective functions. While the choice of the preferred approach is very much dependent on the seriousness attached to the short-term predictions of theory as against the long-run, the use of any approach should be based on the desire to develop a model which has transparent theoretical foundations, and can fit into the historical time series of the data very well (Garratt *et al* (2000)). The modeling approach advocated under this methodology is based on log-linear VAR model estimated subject to long-run relationships obtained from economic theory.

Based on the assumption that individual macroeconomic series has a unit root, each of the long-run relationship derived from the theory should be associated with a cointegrating relationship between the variables, and the existence of such relationships will impose the restrictions on the variables of the model. The use of VAR requires that great care be taken in the initial stages about the choice transformation of data to use to achieve stationarity. VAR approach to long-run modeling in general, has been a source of criticisms in a number of ways. Specifically, Garratt *et al* (2000) suggests three areas of caution when using a VAR modeling technique. First, a VAR model with the first difference of I(1) variables is likely to be mis-specified if there exist a cointegrating relationship between two or more of the I(1) variables (that is, variables integrated of the order of 1). Second, care is needed on the choice of variables to be included in the VAR analysis. Third, where the impulse response function cannot be interpreted with recourse to economic theory, the estimating model will give only a few insights into the economic system that it represents. In order to increase the precision of forecasts based on VARs, Litterman (1986) suggested the combination of unrestricted VAR with Bayesian (This is otherwise referred to as Minnesota priors) analysis.

The structural VAR approach attempts to identify the impulse responses by imposing a priori restrictions on the covariance matrix of the structural errors and/or the long-run responses themselves (Garratt *et al*, 2003). In contrast with the unrestricted VAR approach, structural VARs attempt to provide some economic rationale for the use of covariance restrictions (Such restrictions may pose identification problem for the long-run relationship among variables).

Vector Error Correction Mechanism

The Vector Error Correction (VEC) model is a convenient alternative to the simple VAR model when variables are cointegrated, and provides easy interpretation and differentiation between the short and the long-run implications of the model (Jacobs and Wallis, 2010). It is a restricted VAR

which requires cointegration of some variables. In order to embody the long-run relation of a model, the variables used in the empirical analysis are expected to be I(1) such that it can be incorporated within the dynamic model. Essentially, the VEC model provides a convenient alternative form to the VAR model when variables are cointegrated, and provides easy interpretation and differentiation between the short-run and long-run implications of the model by incorporating into it, an error term which corrects for the adjustment between the short and the long-run behavior of variables.

There is usually some ambiguity over the order of integration of some nominal and short-run variables as they are being transformed or when derived from a system of equations. For instance, an application of Augmented Dickey-Fuller (ADF) test to three sets of price variables such as Δp_t, $\Delta \check{p}_t$, Δp^*_t, representing domestic, foreign and equilibrium interest rates may yield mixed results. In a situation where Phillips and Perron (PP) test cannot prove otherwise, the issue of disagreement is raised about economic modeling and macroeconomic modeling. Following from the validity of Fisher's equation, inflation and interest rate are expected to have the same order of integration (Since changes in both variables arise from the same magnitudes). The theoretical literature generally assumes these variables as I(0)s, but empirical evidence have shown mixed results where interest rate may be behaving as I(1), but with inflation integrated on a different order (Garratt *et al*, 2003). Sometimes in the test for exogeneity in long-run models, as an alternative, the GLPS (An acronym for Garatt, Lee, Pesaran and Shin (2000, 2003 and 2006) model becomes necessary. The GLPS incorporates long-run structural relationships suggested by economic theory as the cointegrating relations of a VECM (Jacobs and Wallis, 2010).

Artificial Intelligence Method

Amiri *et al.* (2011) had faulted the linear functional specifications which were based on OLS regresses. In reaction, they made use of a time series with artificial intelligence method (AIM) to test for the existence of nonlinear relationship among economic variables. The study observed that improved nonlinear Augmented Engel-Granger and Vector Error Correction methods significantly have a better ability to identify long-run cointegration and causal relationships than ordinary linear ones.

Matters Arising

A convenient starting point is to ask if 'long-term' is synonymous with 'long-run'. Frequently, the word, 'long-term' is used in the literature – growth economics, development economics, development planning and financial economics mostly, in a way that could be confused with the long-run. Long-term could be seen in terms of a finite horizon or foreseeable future. Thus, series such as long-term interest rates refer to such rates stipulated for future periods and arrived at after taking into consideration the inflation expected at the relevant future periods. Except for countries operating on the (long-run) natural rate level of output, such future interest rates may not constitute long-run values. Therefore, some care is needed in dealing with such long-term series in long-run analysis. The Lucas critique is also relevant here as such series are usually projected on the basis of ex post data which may not accommodate unexpected future developments even over the horizon being considered.

The second important commentary on this issue is that, policies generating data series following an economic reform may, as recognized by Lucas (1988), be inadvertently admitted as long-run growth variables. In this sense, it is hard to see how some of the policy issues identified in some endogenous growth models that is, apart from technological progress, qualify to be recognized as long-run variables. The effect of expenditure on the national system of innovations boosting capability and revolutionizing technical process continuously cannot be placed on the same pedestal with the provision of infrastructural services, maintenance of law and order, and, regulations of financial markets in long-run growth considerations (Regulations of international trade and protection of intellectual properties are excluded from the list because of the possibility that the former could facilitate greater absorption of new technologies by a follower country and the latter as recognized by Segerstrom (1998), could supply the incentive to innovate and produce new technology). Policies on some of these other issues tend to remain unchanged for a long time and would most likely generate level effects (Adequate provision of infrastructural services and effective law and order are features of development and may be taken for granted in an advanced economy; therefore, they are not likely to be sources of new growth in such an economy). Tagging them 'exogenous' may be insufficient so long as continuous changes are not identified with them. In this wise, studies incorporating these variables into long-run analysis may be in error (Those studies favoring the narrative measures (index) of monetary policy or the financial sector fall into this category. For such studies, a cautionary note exists in Leeper, Sims and Zha (1996) – cited in Walsh (2003) – : '….. most movements in monetary policy instruments represent responses to the state of the economy, not exogenous policy shifts).

A second observation on current usage and application of 'long-run' when posed as a question is, shouldn't there be a distinction between the run-up to the long-run and the long-run itself? The long-run effect in itself could be viewed as two folds viz: in terms of sustaining steady state growth and in terms of shifting the long-run aggregate supply curve. The first basically describes short run economic activities/policies at the steady state, while the second implies long-run shift factors. For countries traditionally viewed as operating on the natural rate level, only activities at the long-run would be of interest. Hence, short term stabilization policies could exist side by side with long-run growth policies capable of shifting outward the production possibilities frontiers. For developing countries usually seen as operating below full employment equilibrium, the focus has mostly been the short run, relying on short term macroeconomic policies to generate growth and higher level of employment.

Lately however, there appears to be a realization of the importance of long-run growth drivers in the policy making circles of these (developing) countries. Hence, these economies put much emphasis (sometimes, rhetorically though) on human capital development and technological innovation. However, the point in this section is that, in the literature, some policies that only contribute to the drive-up to long-run are often specified as long-run variables. Perhaps, the source of the confusion is that, we usually associate long-run with the notion of equilibrium without paying much attention to the differences in stages of growth requirements as noted in opening part of this section. And, as indicated in the preceding paragraph, most innovations in the financial sector fall into the class of short run growth determinants. For example, real interest rate which is often the variable of reference when discussing the benefits of financial liberalization contributes significantly to short run growth but, if we understood the notion of long-run properly, does not

exist at that point, that is, long-run. Most demand for money functions specifying price variables as long-run factors, are also in this error bracket.

A third area of concern in the current usage and application of the notion of long-run in the literature is the extent to which reliance should be placed on nominal variables. Theoretically, all nominal variables generate long-run effect on each other in the sense of run-up to the steady state. To qualify as a possible long-run growth driver, an additional quality is needed and this is exogeneity. It guarantees that such a variable could experience changes even though the economy is in stable long-run equilibrium.

Another pertinent question to ask on the issue of long-run is, what should be the appropriate definition of long-run? The present identification of the concept with the notion of equilibrium appears to reflect an economy-wide or macroeconomy context only. A simple illustration would reveal the inadequacy of this position. Here are two propositions for comparison. 1. Nigeria would be a developed country. 2. Sustained inflation is a monetary phenomenon. Both are long-run propositions, yet, one (as a process) is faster than the other. In effect, every long-run economic process has its own dynamics which is different from those of other processes. Thus, "for some processes, the long-run can be short, for some others, it can be very long (This expression dates back sometime in the literature but the actual publication, I could neither locate nor recollect. In a sense however, Hirshleifer (1980) identifies with this expression)."

In the context of the unit root – cointegration revolution in time series analysis and long-run investigation, the observation of Pesaran (1997) referenced in the preceding section is relevant and deserves some comments. It appears that too often, we forget that the unit root – cointegration development is part of a methodology referred to as 'General to Specific'. Simply described, this methodology helps to identify a congruent model that is both theory consistent and data admissible and also encompasses rival models. Thus, it warrants that econometric models of long-run should be specified according to economic theory, ensuring data admissibility through the investigation of the time series properties (for logical sequence, testing the theoretically identified long-run series for cointegration with the dependent variable if both are I(1) series), obtaining the parsimonious equation and interpreting the results theoretically (In effect, the general to specific methodology usually begins with a large model which is gradually reduced to the parsimonious through an elaborate iterative process. Accordingly, such dynamic single equation models allow for sufficient and varying interactions between the dependent and explanatory variables; they also recognize and provide for the distinction between the effects of endogenous and exogenous variables through a battery of diagnostic tests including that for exogeneity. In this wise, tests for exogeneity in dynamic models (not the cointegration – error correction type) include the Sims – Granger causality. Clearly, the criticisms of such models in Kibritcioglu and Dibroglu (2001) on failure to recognize the possible endogeneity of explanatory variables may not apply to this case). It appears that the failure to realize and follow this basic principle is the source of the myriads of long-run specifications and estimations in the literature. An adaptation of Monetarists' criticism of the structural model evidence of Keynesians is relevant here. Thus, any econometric model of the long-run is only as good as the underlying economic theory.

An analogy to the consideration of the status of nominal variables in long-run analysis is whether structural variables are always of long-run nature. Ordinarily, because of their link with the structure of the economy which can only be altered in a long-run context, there is always that

tendency to regard such variables as of long-run nature. However, a misleading and distorted picture could emerge if the actual context or process is not properly understood while embarking on the model specification. Consider for example, the case of a real export growth function in which the gross domestic product (GDP) is specified as long-run variable in the sense of reflecting the situation of slow technical change. This could go unquestioned depending on the understanding of the analyst and his/er audience or readers. For the better informed, if the relevant economy whose export growth is being modeled is opened, then, the argument changes. The problem of slow technical change can be overcome in the short run through importation. In other words, depending on the composition of imports, a sufficiently open economy can overcome any structural impediment in the short run through importation and with the country putting in place policies (e.g. skill acquisition) necessary to ensure the sustainability of long-run structural adjustment, it is clear under this scenario that, structural factors are best treated as short run determinants (The empirical evidence to support this is the case of the Austrian economy of the 1970s and the 1980s which experienced tremendous expansion in her textile export industry through an import-led innovation strategy. In Africa, the case of Mauritius remains a special study on textile export success).

Related to the above is the need for us to allow economic or socio-economic developments a role in long-run modeling. This is best illustrated with an example. At the onset of a scourge e.g. HIV/AIDS, with the notice of absence of cure, the rest of the society not only avoid the infected but the productivity of the victims too falls considerably and progressively till the ultimate. Treating the incidence of this malaise as a long-run variable via its human capital growth effect and hence, productivity slow down, appears to be smart economics. With time and breakthrough in the search for medical solution and credible treatment of the ailment coupled with increasing public enlightenment campaign on the need to accommodate and tolerate the affected, people no longer avoid the victims raising their morale. Besides, the spread of the disease is checked through the awareness campaign. And, with the help of proper medication, the productivity impact becomes considerably less. In this situation, to continue specifying the incidence of the scourge as a long-run variable would amount to a case of economics lacking in sophistication.

Two other examples of inadmissible factors in long-run modeling could further throw some light on the relevant issue under discussion. First, is the case of the ever so popular and important productivity growth, which owes its measurement procedure to Professor Robert Solow. Accordingly, it became known as the "Solow residual" which reflects the way it is generated. Quite simply, the idea is to explain observed growth of output beyond the fraction accounted for by factor inputs, mainly, labor and capital. The residual is thus generated from the growth regression involving the factor inputs as explanatory variables. Most humbly, it is submitted that this procedure amounts to finding or searching for, the missing component of an identity and this is not the same as ascertaining the factors accounting for the movements in output. A favored approach in this regard first specifies both the short and long-run determinants of productivity growth which should be computed as the percentage change in the sum of the ratios of the gross national product to capital stock and total labor input respectively. The short-run determinants would include real interest rate, real exchange rate and the supply (that is, stock) of factor inputs while the long-run determinants (fundamentals) would include technological progress (as may be captured variously by expenditure on the national system of innovations, regulations of

international trade and protection of intellectual properties)(In most economies, especially, the developing type, the incidence of corruption could be a fundamental in productivity growth consideration in the sense of hypothetically causing the full information natural rate level of output to fall short of the full capacity output). The fitted series of the regression involving the dependent variable and the fundamentals should be a more credible long-run growth determinant (shift factor).

Finally, consider a balance of payments function (BOP) into which imports, exports, interest rate, price level, exchange rate or real exchange rate, real money supply and real gross national product are specified as long-run determinants. Quite clearly, interest rate is a measure of cost of capital (not real cost though) in this model while price level changes track inflation effect. However, the two variables, interest rate and price level, generate opposing effects on the dependent variable and are also expected to grow proportionally relative to each other in the long-run. The overall (that is, long-run) effect of the inclusion of the two variables in the model is zero. Besides, they both would change equip-proportionally in the long-run relative to a change in the exchange rate. Both imports and exports simply are symptoms of disequilibrium in the BOP and not fundamental causes of such disequilibrium. Real money supply would technically have no long-run effect as there is full offsetting effect of price level on money supply changes. For most economies, nominal or real exchange rate would be stable in the long-run. Clearly, in this scenario, only the real GNP qualifies as the long-run fundamental in the sense of its growth generating an adverse development in the BOP. This clearly suggests the need for careful thought in specifying the long-run variables for any economic process.

CONCLUDING REMARKS

This survey and the issues generated were designed to assist the developing professional economists to overcome the confusion that is likely to be created by the numerous usages (implicitly therefore, definitions) and applications of the notion of long-run in the literature. This paper therefore should be seen as just a supplement in the learning process.

The various approaches to long-run analysis in the growth, financial and econometric branches of economics were identified and presented briefly in a non-technical manner. Observations on current practices in these areas were made and suggestions as to how to overcome the related pitfalls were proffered. These suggestions were wide-ranging, covering, matters of working definition, variables' admissibility, status of nominal and real variables, economic theory as the fulcrum of long-run econometric designs, the influence of dynamics in specifications as well as the computation of a notable long-run growth factor. As practicable, instances were used to illustrate the various viewpoints. These experiments came mostly from this writer's research endeavors as well as those conducted by some of his graduate students.

It is of utmost importance for the learner to always remember that, it is not compulsory always to undertake long-run analysis. Most stabilization policies (monetary or fiscal) target the short run. Hence, searching for long-run in respect thereof may be a futile exercise. No matter how well done, a long-run analysis that is unnecessary simply interprets as 'doing in a beautiful or fanciful way, what ought not to have been done at all'. It is meaningless, useless and a waste of

time and effort. Where it is absolutely desired and deemed necessary to conduct long-run analysis, careful thoughts should be given to model specification and the subsequent analysis.

REFERENCES

Afonso, A., P. Gomes, and P. Rother 2007. Short and long-run determinants of sovereign debt credit rating.Retrieved from http://economics.ca/2008/papers/0006.pdf

Aghion, P. and P. Howith 1992. A model of growth through creative destruction. Econometrica 60: 323 – 351.

Aghion, P. and P. Howith 1992.1994. Endogenous technical change: the Schumpeter perspective. In Economic Growth and the Structure of Long-Term Development, ed. L.L.Pasinetti and R.M. Solow. New York: Martin's Press.

Aghion, P. and P. Howith 1992.1998, Endogenous Growth Theory. Cambridge Mass: MIT Press.

Al-Azzam, A. and D. Hawdon 1998. Estimating the demand for energy in Jordan: a Stock-Watson dynamic OLS (DOLS) approach. Retrieved from http://ideas.repec.org/p/sur/seedps/97.html

Alonso, A., C. Echevarria and K. Tran 2002. Long-run economic performance and the Labor market. Retrieved from http://www.arts.usask.ca/economics/faculty/papers/aet.pdf

Askari, A. and T. Cummings 1977. Estimating agricultural supply response with the Nerlove model: a survey. *International Economic Review*, 18: 257-329.

Amiri, A., U. Gerdtham, and B. Ventelou 2011. A new approach for estimation of long-run relationships in economic analysis using Engel-Granger and Artificial Intelligence methods," retrieved from halsh.archive-ouvertes.fr./docs/00/60/60/48/PDF/PDT/DTGREQAM2011_31.pdf

Arrow, K. 1962. The economic consequences of learning by doing. *Review of Economic Studies*29(80): 155 – 173.

Banerjee, A., J. Dolado, D.F. Hendry and G. Smith 1986. Exploring equilibrium relationships in econometrics through static models. *Oxford Bulletin of Economics and Statistics*, 48: 253-77.

Barro, R.J.1991. Economic growth in a cross section of countries. *The Quarterly Journal of Economics* 106(2): 407 – 443.

Barro, R.J. and G.S. Becker1989. Fertility choice in a model of economic growth. *Econometrica* 57: 481 – 501.

Barro, R.J.and X. Sala-i-Martin (2004). Economic Growth, 2nd edition. Boston: MIT Press.

Becker, G.S., M. Kevin and R. Tamura 1990. Human capital, fertility and economic growth. *Journal of Political Economy* 98: S12 – 37.

Bencivenga, V.R. and B.D. Smith 1991. Financial intermediation and endogenous growth. *Review of Economic Studies* 58: 195 – 209.

Blough, S. R. 1988. On the impossibility of testing for unit roots and cointegration in finite samples.*Working Paper*, No. 211, Department of Economics, Johns Hopkins University.

Boschen, J.F. and L.O. Mills 1991. The effects of countercyclical monetary policy on money and interest rates: an evaluation of evidence from FOMC documents. *Working Paper 91 – 20, Federal Reserve Bank of Philadelpia*, October.

Brayton, F. and E. Mauskopt1985. The Federal Reserve Board MPS quarterly econometric model of the U.S. economy. *Economic Modeling* 2:170-292.

Brayton, F. and P. Tinsley 1996. A guide to FRB/US: a macroeconomic model of the United States. *Finance and Economics Discussion Series* 42, Washington, D.C.: Federal Reserve Board.

Buffie, E.F. 1984. Financial repression, the new structuralists and stabilization policy in semi-industrialized economies. *Journal of Development Economics* 14(3): 305 –322.

Canning, D., and P. Pedroni 1999. Infrastructure and long-run economic growth. Retrieved from http://economics.cornell.edu/cae/infrastructure-7-99.pdf

Cass, D. 1965. Optimum growth in an aggregative model of capital accumulation. *Review of Economic Studies* 32: 233 – 240.

Cass, D.1972. On capital over-accumulation in the aggregative, neoclassical model of economic growth: a complete characterization. *Journal of Economic Theory* 4: 200 – 223.

Cho, Y. J. 1986. Inefficiencies from financial liberalization in the absence of well-functioning equity markets. *Journal of Money, Credit and Banking* 18: 191 – 199.

Cho, Y.J. 1988. The effects of financial liberalization on the efficiency of credit allocation: some evidence from Korea. *Journal of Development Economics* 29: 101 – 110.

De Gregorio, J. 1992. The effects of inflation on economic growth: lessons from Latin America. *European Economic Review* 36: 417 – 425.

Dees, S., F. di Mauro, M.H. Pesaran and L.V. Smith 2007. Exploring the international linkages of the euro area: a global VAR analysis. *Journal of Applied Econometrics*, 22: 1–38.

DeLong, J. B. 2001. Long-run economic growth. Retrieved from http://econ161.berkeley.edu/

Dinopoulos, E. and P. Thompson 1999. Scale effects in schumpeterian models of growth. *Journal of Evolutionary Economics* 9: 157 – 186.

Domar, E.D. 1946. Capital expansion, rate of growth and employment. *Econometrica* 14: 137 – 147.

Dritsakis, N. 2004. Tourism as a long-run growth factor: an empirical investigation for Greece using causality analysis. Retrieved from http://users.uom.gr/~drits/publications/tourism-economic_growth.pdf

Emran, M. S. and F. Shilpi 2008. Estimation import demand function in developing countries: a structural econometric approach with application to India and Sri Lanka. Institute of International Economic Policy *Working Paper Series*, George Washington University.

Engle, R.F. and C.W.J. Granger 1987. Cointegration and error correction: representation, estimation and testing. Econometrica 55: 251-76.

Friedman, M. 1968. The role of monetary policy. *American Economic Review* 58: 1 – 17.

Fry, M.J. 1978. Money and capital or financial deepening in economic development. *Journal of Money, Credit and Banking* 10: 465 – 475.

Fry, M.J. 1980a. Saving, investment, growth and the cost of financial repression. *World Development* 8: 317 – 327.

Fry, M.J. 1980b. Money, interest, inflation and growth in Turkey. *Journal of Monetary Economics* 6: 535 – 545.

Fry, M.J. 1981. Models of financially repressed developing economies. *World Development* 10: 1049 -1057.

Galbis, V. 1977. Financial intermediation and economic growth in less developed countries: a theoretical approach. *Journal of Development Studies* 13(2): 58 – 72.

Grossman, G.M. and E. Helpman 1991. Quality ladders in the theory of economic growth. *Review of Economic Studies* 58(1): 43 -61.

Garratt, A., K. Lee, M.H. Pesaran and Y. Shin 2000. A structural cointergating VAR approach to macroeconomic modeling. In Econometric Modeling: Techniques and Applications, ed. S. Holly, M.R. Weale,94-131. Cambridge MA: Cambridge University Press.

Garratt, A., K. Lee, M.H. Pesaran and Y. Shin2003. A Long-run structural macroeconomic model of UK. *The Economic Journal* 113(487):412-455.

Garratt, A., K. Lee, M.H. Pesaran and Y. Shin 2005. Global and National Macroeconometric Modeling*: A* Long Run Structural Approach. Oxford: Oxford University Press.

Granger, C.W.J. 1981. Some Properties of Time Series Data and their use in Econometric Model Specification. *Journal of Econometrics*, 16:121-130.

Gujarati, D.N. 2004. Basic Econometrics, 4th ed. Tata McGraw-Hill Publishing Company Limited, New Delhi.

Harrods, R.F. 1939. An essay in dynamic theory. *Economic Journal* 49: 14 – 33.

Jacobs, J. P. A. M., and K.F. Wallis 2010. Cointegration, long-run structural modeling and weak exogeneity: two models of the UK economy. *Journal of Econometrics,* 158: 108-116.

Johansen, S. 1988. Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 12: 231-254.

Johansen, S. 1991. Estimation and hypothesis testing of cointegration vectors in Guassian vector autoregressive models. *Econometrica*, 59: 1551-80.

Jones, C. 1995. R and D based models of economic growth. *Journal of Political Economy* 103(4): 759 – 784.

Jones, L. and R. Manuelli 1990. A convex model of equilibrium growth: theory and policy implications. *Journal of Political Economy*, 98: 1008-1038.

Kapur, B.K. 1976. Alternative stabilization policies for less developed economies. *Journal of Political Economy* 84(4): 777 – 795.

Kibritcioglu, A. and S. Dibooglu 2001. Long-run economic growth: an interdisciplinary approach.Retrieved from http://www.business.uiuc.edu/Working_Papers/papers/01-0121.pdf

Kim, J. H., I. Fraser and R.J. Hyndman 2007. Improved interval estimation of long-run response from a dynamic linear model: a high density region approach. Retrieved from http://www.robjhyndman.com/papers/longrun.pdf

King, R.G. and R. Levine 1993. Finance, entrepreneurship and growth: theory and evidence. *Journal of Monetary Economics* 108: 513 – 542.

Knight, F. 1984. Diminishing returns from investment. *Journal of Political Economy* 52: 26 -47.

Koopmans, T.C. 1965. On the Concept of Optimal Economic Growth. New Haven: Yale University.

Kortum, S. 1997. Research, patenting and technological change. *Econometrica* 65: 1389 – 1420.

Leeper, E.M., C.A. Sims and T. Zha1996. What monetary policy do? *Brookings Papers on Economic Activity 2: 1 – 63.*

Li, H. and G.S. Maddala 1999. Bootstrap variance estimation of non-linear function of parameters: an application of long-run elasticities of energy demand. *Review of Economics and Statistics* 81: 728 – 733.

Littterman, R. 1986. Forecasting with Bayesian autoregressions – five years of experience. *Journal of Business and Economic Statistics* 4:25-38.

Lucas, R. E. Jr. 1976. Econometric policy evaluation: a critique. In The Phillips Curve and Labor Markets, ed.K. Brunner and A. Meltzer. *Carnegie Rochester Series on Public Policy*, 1: 19-46.

Lucas, R.E. Jr. 1988. On the mechanics of economic development. *Journal of Monetary Economics*, 22: 3-42.

Macquez, J. and C. McNeilly 1998. Income and price elasticities for exports of developing countries. *The Review of Economics and Statistics*, 70: 306-314.

Marin, D. 1988. Import-led innovation: the case of the Austrian textile industry. *WeltwirtshaflichesArchiv* 124: 550 – 565.

Mathieson, D. 1980. Financial reform and stabilization policy in a developing economy. *Journal of Development Economics* 7: 359 – 395.

Mckinnon, R.I. 1973. Money and Capital in Economic Development. Washington, D.C.: The Brookings Institution.

Miller, R.L. and D.D. Vanhoose 1993. Modern Money and Banking, 3rd edition. New York & co.: McGraw-Hill Inc.

Monfort, B. and C. Mulder 2000. Using credit ratings for capital requirements on lending to emerging market economies- possible impact of the new Basel accord. Washington, D.C.: International Monetary Fund, *Working Papers* 00/69.

Montiel, P. 2003. Macroeconomics in Emerging Markets. Boston: Cambridge University Press.

Mora, N. 2006. Sovereign credit ratings: Guilty beyond reasonable doubt. *Journal of Banking and Finance*, 30: 2041-2062.

Murphy, C.W. 1988. An overview of the Murphy model. *Australian Economic Papers* 27:175-99.

Ogun, O. 2004. Real exchange behavior and non-oil export growth in Nigeria. *African Journal of Economic Policy* 11: 69 – 89.

Ogun, O. 2008. Purchasing power parity exchange rate, misalignment and export growth in Nigeria. *Ibadan Journal of the Social Sciences* 6: 79 – 85.

Ogun, O. 2012. Inflation fundamentals and growth in a developing economy. *The American Charter of Economics and Finance* 1: 23 – 59.

Ogun, O. 2012. An exposition on the demand for money. Processed.

Patibandla, M. 1996. International trade and long-term economic growth: a few Issues of growth strategy for India. *Economic and Political Weekly*, 31 : 3121-23, 3125-28.

Peretto, P. 1998. Technological change and population growth. *Journal of Economic Growth* 3: 283 – 311.

Peretto, P. and S. Smulders 2002. Technological distance, growth and scale effects. *Economic Journal* 112: 603 – 624.

Pesaran, M. H. 1997. The role of economic theory in modeling the long-run. *The Economic Journal*, 107: 178-191.

Pesaran, M.H. and Y. Shin 1996. An autodistributive lag modeling approach to cointegration analysis. *DAE Working Paper* no. 9514, Department of Applied Economics, University of Cambridge.

Pesaran, M. H., T. Schuermann and S.M. Weiner 2004. Modeling regional interdependencies using a global error-correcting macroeconometric model. *Journal of Business and Economic Statistics*22: 129–181.

Ramsey, F. 1928. A mathematical theory of saving. *Economic Journal* 38: 543 – 559.

Rebelo, S.T. 1991. Long-run policy analysis and long-run growth. *Journal of Political Economy*, 99: 500-521.
Ricardo, D. and G. Eck 1817. The Principles of Political Economy and Taxation. London: Prometheus Books.
Romer, C.D. and D.H. Romer 1989. Does monetary policy matter? a new test in the spirit of Friedman and Schwartz .In *NBER Macroeconomics Annual 1989*, eds. O.J. Blanchard and S. Fischer, 121 – 170, Cambridge MA: MIT Press.
Romer, P.M. 1986. Increasing Returns and Long-run Growth. *Journal of Political Economy* 94: 1002-1037.
Romer, P.M. 1987. Growth based on increasing returns due to specialization. *American Economic Review* 77: 56-62.
Romer, P.M. 1989a. Human capital and growth: theory and evidence. Washington, D.C.: *National Bureau of Economic Research* WP 3173
Romer, P.M. 1989b. What determines the rate of change of technological change? Washington, D.C.: *World Bank PPR WP*279.
Romer, P.M. 1989c. Capital accumulation in the theory of long-run growth. In Modern Business Cycle Theory, ed. R. Barro, 51 – 127. Cambridge MA: Harvard University Press.
Romer, P.M.1990. Endogenous technological change .*Journal of Political Economy* 98: 71-102.
Roubini, N. and X. Sala-i-Martin 1992. Financial repression and economic growth. *Journal of Development Economics* 39: 5–30.
Sachs, J. and A. Warner 1995. Economic reform and the progress of global integration. *Harvard Institute of Economic Research Working Paper* 1733.http://EconPapers.repec.org/RePEC: fth:harver: 1733
Schumpeter, J. 1934. The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest and theBusiness Cycle. Boston: Harvard University Press.
Seddighi, H. R., K. A. Lawler and A. V. Katos 2000. Econometrics: A Practical Approach. London and New York: Routledge, Taylor and Francis Group.
Segerstrom, P. S. 2000. The long –run growth effects of R&D subsidies. Retrieved from http://www. economics.harvard.edu/faculty/kremer/files/Economic%20transformation_080918%20To%20Post.pdf
Sheshinski, E. 1967. Optimal accumulation with learning by doing.In*Essays on the Theory of Optimal Economic Growth,* ed. K. Shell.Cambridge Mass: MIT Press.
Sims, C.A. 1972. Money, income and causality. *American Economic Review* 62: 540 – 542.
Sims, C.A. 1980a. Comparison of interwar and postwar business cycles. *American Economic Review*70: 250 – 257.
Sims, C.A.1980b. Macroeconomics and reality. *Econometrica* 48: 1-48.
Smith, A. 2003. The Wealth of Nations, 3rd edition. New York: Bantam Dell.
Solow, R. 1956. A contribution to the theory of economic growth. *Quarterly Journal of Economics* 70: 65 – 94.
Solow, R. 1957. Technical change and the aggregate production function. *Review of Economics and Statistics 39: 312-320.*
Stock, J.H. 1987. Asymptotic properties of least squares estimators of cointegrating vectors. *Econometrica* 56: 1035-56.
Swan, T.W. 1956. Economic growth and capital accumulation. *Economic Record* 32: 344 – 361.
Taylor, L. 1983. *Structuralist macroeconomics: applicable models for the third world.* New York: Basic Books.
vanWijnbergen, S. 1983a. Interest rate management in LDCs. *Journal of Monetary Economics* 12: 433 – 452.
vanWijnbergen, S. 1983b. Credit policy, inflation and growth in a financially repressed economy. *Journal of Development Economics* 13: 45 – 65.
Utkulu, U. 1994. How to estimate long-run relationships in economics: an overview of recent developments. retrieved from http:/click.infospace.com/ClickHandler.ashx?du=http%3a2fkisi.deu.edu.tr%2futku.utkulu%2fdosyalar%2fHow_to_estimate.DOC&ru=http%3a%2f%2fkisi.deu.edu.tr%2fdosyalar%2FHOW_to_estimate.DOC&ld=20120201&ap=1&app=1&c=mammaxml2&s=mammaxml2&coi=230137&cop=maintitle&ep=1&cuip=41.206.12.16&npp=1&p=0&pvaid=abf949715e49399ee3b4f59b6ade5&hash=E22375EAD6D5FDCF980196FOFFF57278
Wallich, H. 1969. Money and growth: a country cross section. *Journal of Money, Credit and Banking* 1: 281 – 302.
Wallis, K.F. and J.D. Whitely 1987. Long-run properties of large-scale macroeconometric models. *Annals of Economics and Statistic* 6.: 207-224.
Walsh, C.E. 2003. Monetary Theory and Policy. Cambridge Mass.: MIT Press.

Wangwe, S.M. 1995. Exporting Africa: Technology, Trade and Industrialization in Sub-Saharan Africa. UNU/INTECH Studies in New Technology and Development. London and New York: Routledge.
Young, A. 1928. Increasing returns and economic progress. *Economic Journal* 38: 527 – 542.
Young, A. 1998. Growth without scale effects. *Journal of Political Economy* 106: 41 -63.

THE BUY LOCAL INITIATIVE AND ITS EFFECTIVENESS IN A SMALL ISLAND ECONOMY: EVIDENCE FROM THE PACIFIC ISLAND OF GUAM

Maria Claret M. Ruane, University of Guam

ABSTRACT

Small islands around the world are constantly facing the challenge of how to enhance the growth of their economies so as to provide an increased standard of living for their population. The major culprit is these islands' "smallness", which reflects itself in different ways (small population, low total income and per capita income, limited productive resources) results in a lack of economies of scale and economic diversification, which limits opportunities for local production and for producing for exports and instead increases reliance on imports for a wide array of goods and services. Efforts to address these challenges look at solutions that specifically focus on the local economy by trying to keep income and spending from leaking out and instead multiplying within the local economy. In Guam, these efforts lead to initiatives that encourage local buying (as both a short-run and long-run solution), which is complemented by local hiring and local producing as long-run solutions. This paper discusses this initiative in Guam and evaluates its effectiveness in making a positive contribution to Guam's economic development.

INTRODUCTION

"Buying local" can truly be a positive contributor to an economy, especially for a small island economy like Guam, a territory of the United States located in the Western Pacific Region. The resulting positive contribution to the local economy increases with the amount that local consumers buy from local businesses, since it is these businesses that provide the majority of local employment, pay taxes to the local government and give back to our larger island community. The University of Guam-Pacific Center for Economic Initiatives director, Dr. Anita Borja Enriquez, makes this point (*Marianas Variety*, October 20, 2011):

> *"you make the decision to 'Buy Local' when you buy products that are sold locally instead of ordering them 100 percent from off-island; by hiring Guam residents first before off-island hires; and by using Guam-based businesses instead of 'outsourcing work' off-island."*

THE BUY LOCAL INITIATIVE IN GUAM

The Buy Local initiative in Guam, also known as the "Local Guam Movement… Think, Support, Buy", was established in 2010 as a collaborative effort by the University of Guam-Pacific Center for Economic Initiatives, Guam Chamber of Commerce (including the Guam Young Professionals) and the Guam Economic Development Authority (http://www.buylocalguam.org). It …

"is a targeted marketing education campaign that promotes the economic benefits of supporting Guam-based businesses and organizations and Guam-based hiring. It is intended to motivate Guam consumers, businesses, and other organizations to keep the dollars circulating in Guam and shift spending by at least 10% towards Guam-based businesses. It will also promote providing more jobs for Guam's residents. The resulting benefits are more dollars for our local tax base to support public health care, safety, and education; roads and water systems; public parks; reduced carbon-footprint; and better balance of trade. It aims to motivate Guam consumers, businesses and other organizations to keep the dollars circulating in Guam and shift spending by at least 10 percent to Guam-based businesses."

Since then, events and activities have taken place to educate local residents about the idea behind the Buy Local initiative and to promote the initiative (see examples in Appendix A). Local companies and organizations worked with University of Guam-Pacific Center for Economic Initiative, Guam Chamber of Commerce, Guam Young Professionals, Guam Economic Development Authority, Guam Preservation Trust, Hagåtña Restoration and Redevelopment Authority of the Department of Chamorro Affairs, Guam Humanities Council and *Local First! Guam* to sponsor many of these Buy Local activities and events.

ECONOMIC BENEFITS FROM A BUY LOCAL INITIATIVE

A common way to measure how much local spending stays and circulates (or multiplies) in the local economy is to use the concept of spending multiplier. The spending multiplier captures the process of spending an additional $1 "in the first round" and how this $1 leads to subsequent rounds of spending in that, each additional $1 spent in the local economy becomes income to some local businesses and their employees, from which they would spend in the next round, thus becoming income to some other local business and their employees, from which they would spend in the next round, and so on. In the end, each additional dollar spent generates an amount greater than the dollar spent "in the first round". Hence, each additional dollar spent "in the first round" is multiplied. If we know what the spending multiplier is that applies to Guam's economy, we could estimate the resulting total spending on the local economy and then extrapolate the implications for additional tax revenues from Gross Receipts Tax and additional jobs created.

The Buy Local initiative website (http://www.buylocalguam.org) lists the following as *Top Reasons to Think, Support and Buy Local*:

- *More money for Guam. When you buy from an independent, locally owned business, it is more likely that they will then purchase from other local businesses and service providers, thereby strengthening the economic base and keeping the dollars circulating on island.*
- *Job creation. Hiring Guam residents first creates and maintains much-needed jobs to raise the standard of living and support the overall quality of life of Guam residents.*
- *Big impact. Consider shifting 10% of the purchases you usually make from the "Three O's" (on-line, on-base, and/or off-island) to local businesses that provide the same products at reasonable prices. This small change will make a big difference in the local economy.*
- *Support community groups. Small, local business owners are more likely to support Guam-based, non-profit organizations and events in the community.*
- *Reduce environmental impact. Locally owned businesses make more local purchases, thereby requiring less transportation. This generally means contributing less to congestion and environmental pollution.*

- *Invest in the community. Local businesses are owned by people who live in the community, are less likely to leave, and invest more in Guam's future.*
- *Build Guam's tax revenue base. Local businesses would directly contribute to Guam's tax base, adding much-needed revenue to support public health, public safety, public roads, water, sewer systems, utilities, public education, community projects, and other public service initiatives.*

Ruane (2011) provides a theoretical framework for the Buy Local initiative, starting with a standard version of the spending multiplier is given as

$$\text{Standard Spending Multiplier} = 1 / (1-MPC(1-t) + MPIm) \quad (1)$$

Where MPC=marginal propensity by households to consume out of their total/gross household income; t = marginal tax rate on consumer/households' incomes and MPIm = marginal propensity to import. Formula (1) shows the spending multiplier increases with higher MPC, lower tax rate and lower MPIm. These three factors correspond to the three leakages (using John Maynard Keynes' term) from households' ability to spend on the local economy: household saving, taxes, and imports.

In Guam's case, the standard spending multiplier needs to be modified in order to reflect additional leakages from the Guam's economy when household spending goes toward spending at stores in military bases located in Guam, from online sources and from off-island sources. Elsewhere (in Ruane, 2011), I explained that only a portion a_1 of the MPC goes to the local economy as summarized in Formula (2):

$$\text{Guam's spending multiplier} = 1 /(1- a_1MPC(1-t) + MPIm)) \quad (2)$$

i.e., only a proportion a_1 of household spending are multiplied in the local economy. The balance $(1-a_1)$ is assumed to have leaked out. This is in addition to the amount of spending that leaked out through imports.

The policy implications for increasing Guam's spending multiplier are straightforward and as easy as referring to Formula (2). First is to find ways to increase the marginal propensity to consume (MPC). However, caution must be taken in making the distinction between short-run and long-run benefits to the economy and the balance between consumption and saving. Second is to reduce the marginal tax rate, which is not an option for Guam since it adopts the marginal tax rates (t) that prevail in the Continental U.S. Third is to find ways to switch local spending toward local sellers, as opposed to military base, online sales, off-island shopping and imports. If this sounds familiar, this is what forms the basis for the Buy Local initiative. Lastly, one sure way to increase Guam's spending multiplier is to reduce its current dependence on imports by perhaps developing industries that would produce local goods that would substitute for those currently imported, in this case, a "Local Producing"initiative. As Guam Chamber of Commerce President David Leddy said, "development of local industries goes hand in hand with the Buy Local campaign". (*Marianas Business Journal*, 2012, July 16)

Unfortunately, we are not sure what the spending multiplier is for Guam. Estimates by local economists range from 1.2 to 2.0. In December 2011, a study estimated the local spending

multiplier to be 1.30 (*University of Guam-Pacific Center for Economic Initiatives Technical Report, Issue 15*) based on the following calculation:

$$\text{Guam's spending multiplier} = 1 /(1- 0.71 + 0.48)) = 1.2987(\text{rounded off to } 1.30)$$

This 1.30 estimate already took into account spending/purchasing power that "leaked" out from the local economy to pay purchases from online sources (6.53% of Guam consumers' annual purchases), off-island shopping (6.86%) and military base stores (15.78%). This means that 71% of Guam consumer's annual purchases go to local vendors but keep in mind that the full 71% does not stay in the local economy because local vendors purchase some of their merchandise from imported suppliers. The U.S. Department of Commerce-Bureau of Economic Analysis estimates of Guam's Gross Island Product show that imports make up 48% (BEA, 2012, September 24). The formula and the above explanation imply that the spending multiplier on the local economy can be increased by reducing our purchases from non-local sources, that is, by "buying local."

The above theoretical framework, combined with the numerical estimate for the local spending multiplier, allow us to calculate the potential positive contribution of "buying local" on the Guam economy as follows:

• The spending multiplier of 1.30 says that every additional $100 spent in the local economy would generate an additional $30 on top of the $100 spent in the "first round", or a total of $130 of additional spending/income.

• This additional spending represents additional sales receipts of $130 would result in additional gross receipts taxes of $5.20 (=4% of $130). Also, this additional spending represents additional income of $130, which would generate additional income taxes of $19.50, calculated at an assumed average tax rate of 15%.

• The impact on job creation can be estimated using the following information: Guam's economy is approximately $4 billion per year (using latest estimate for year 2010) and provides around 60,000 jobs (rounding off recent Guam Department of Labor employment data). This suggests that every $100,000 worth of spending/sales/income on the local economy would yield 1.5 local jobs.

SURVEY ON THE EFFECTIVENESS OF THE BUY LOCAL INITIATIVE

Methodology

This survey was developed in early 2012 under the leadership of the University of Guam-Pacific Center for Economic Initiatives Director, Dr. Anita Borja Enriquez, with inputs from Drs. Karri Perez, Annette Taijeron Santos and Fred Schumann and Ms. Eileen Agahan. A shorter version of it was piloted in July at the Chamorro Village in conjunction with the Guam Buy Local initiative called "Think, Support, Buy Local at Chamorro Village", which ran from June 22 to July 18, 2012 (http://www.buylocalguam.org). A longer version of a related survey was conducted on September 15 at the Micronesia Mall during a UOG-organized event called "Green Energy Career Expo". In November, the survey instrument was expanded by Dr. Claret Ruane in order to measure a change in awareness, perception and support of the initiative, with the goal of measuring (quantifying to the extent possible) the effectiveness of the Buy Local initiative approximately one

year since the marketing campaign was launched. This report will focus on the last survey conducted.

The survey consisted of 13 questions designed to measure Guam consumers' awareness, perception and support of the Buy Local initiative and the initiative's effect on consumers' purchase decisions. The survey instrument received the approval of the Committee on Human Research Subjects as #12-111.

Several survey questions were designed to capture different levels of involvement with the initiative or the idea behind the initiative:

- Awareness is the lowest level of involvement and constitutes having heard of the initiative or having some information/understanding about the idea behind the initiative.
- Perception of importance is the next level of involvement and is a progression from being aware of the initiative to developing an opinion about it in general and its importance, in particular. At this level, one would agree or disagree, to different degrees, that the initiative is important. This includes agreeing that the initiative is important *in principle* but not necessarily in practice.
- Being supportive or giving support is the highest level of involvement among the three considered here and translates awareness of the initiative and the belief in its importance into action of support or moving the initiative forward. This level is a manifestation of *putting the idea/principle behind the initiative into practice.*

The ultimate goal of this survey is to measure in a quantifiable manner the effectiveness of the Buy Local initiative in terms of how much desired action resulted from the initiative, with desired action measured as *additional purchase from local vendors and reduced purchase from non-local vendors*, which include military base stores, online sources and purchases made off-island during a trip and brought back to Guam.

Conducting the survey

The survey was conducted in face-to-face format on November 10 and 17 at Agana Shopping Center as part of Learn Local event aimed at "promoting the importance of thinking, supporting, and buying local" (http://www.buylocalguam.org). The Learn Local event was announced through a variety of means, including the Buy Local website http://www.buylocalguam.org, UOG main website, news coverage by the local media, email to social networks and by word of mouth. Responses to the survey were later inputted in Microsoft Excel.

Survey Respondents

A total of 435 individuals participated in the survey. Efforts were made to achieve a sample of respondents with diversified backgrounds, but because the survey is voluntary, this goal was not always achieved. The respondents' profile is presented in the Appendix B.

SURVEY RESULTS

First Learned About "Buy Local"

When respondents were ask where they first learned about "Buy Local" in Guam, 36.36% heard about it first on the radio, 16.32% from family members, friends and someone else (not including from a teacher in a classroom-setting), and 14.45% from TV. Although respondents were asked to choose only one answer, 10.02% indicated multiple answers, several of them included the same three channels already mentioned (radio, word-of-mouth and TV). The results are found in Table 1.

TABLE 1: Where did you first learn about "Buy Local" in Guam? (Check one only)

Radio	156	36.36%	#1
TV	62	14.45%	#3
Newspaper (Of the 31 respondents, 20 indicated *Pacific Daily News*)	31	7.23%	
Magazine	0	0.00%	
Word-of-mouth	78	18.18%	#2
Teacher/In class	8	1.86%	
Multiple answers	43	10.02%	
Other answers, including *In a business/restaurant/store or at a mall (23 respondents)* *UOG (6 respondents)* *Guam Chamber of Commerce/Guam Young Professionals (2 respondents)* *First time/at the survey booth (10 respondents)*	51	11.89%	
Responses	**429**		
Skipped	**22**		

What "Buy Local" Means

From the beginning of the Buy Local initiative and even while the marketing and education campaign was going, one challenge the initiative faced is the meaning of the phrase "Buy Local". Although the initiative was clear in its use of "Buy Local" as "supporting Guam-based businesses and organizations and Guam-based hiring", implying that the term "local" means "Guam-based" or "located in Guam". (http://www.buylocalguam.org)

Another attempt to clarify the meaning of a "local business" is made in the December 2011 publication entitled *Buy Local Holiday Gift Guide*, which was circulated as an insert in the *Pacific Daily News*. That publication stated that "… local businesses are broken down into three different tiers: locally-owned and independent businesses, businesses that offer local products and services, and businesses that support the local community." Whereas the first explanation emphasized the location of the business (and hiring), i.e., "Guam-based", the latter statement brought up issues such as ownership of the business, product offerings as well as support for the local community. These two interpretations of the term "local" highlight why we expect local residents and survey respondents to have different interpretations or understandings of the meaning of the term "local" or "buy local". This concern motivated the survey question "What do you think "Buy Local"

means?" Respondents were given six different possible answers or were given the option to write in their own interpretations. The most popular answer indicated by 41.12% of the respondents interpreted "Buy Local" to mean "Buying products that are made in Guam from Chamorro-owned stores/businesses located in Guam", which highlights the type of product (Guam-made), the ownership of the business (Chamorro) and the location of the business (Guam). Note that "Chamorro" refers to the native people of the island of Guam and its culture. 17.76% interpreted "Buy Local" as "Buying products that are made in Guam from stores/business located in Guam (including military base stores), even they are not Chamorro-owned", i.e., product made in Guam sold by a business located in Guam regardless of the ethnicity of the business owners. The third most popular interpretation of "Buy Local" representing 13.55% of the respondents is "Buying products INCLUDING both made on and outside of Guam from Chamorro-owned stores/businesses located in Guam," which focused on the ownership of the business (Chamorro) and location (Guam) but not where the product was made. A summary of survey answers to this question is in Table 2.

Effectiveness of the Buy Local Initiative

How effective has the Buy Local marketing/educational campaign been after one year? To do this, the survey question tries to measure local residents' awareness of the initiative, how important they perceive of it and how supportive they have been of it. The survey question then asks to what extent the initiative has translated into desired action, in this case, to a shift of purchase in favor of local vendors.

Table 3 shows an increase in awareness of the Buy Local initiative over a one-year period, with awareness measured by respondents indicating that they are "aware", "very aware" or "extremely aware" of the Buy Local initiative one year later (in this case, in November 2012 when the survey was conducted) compared to their awareness one year earlier. Whereas only 50.35% of the respondents were aware (as defined above) of the initiative a year ago, 83.29% are aware of the initiative one year later.

Table 4 also shows an increase in the respondents' perception of how important buying local is to the island economy of Guam, from 80.47% of respondents indicated that they think buying local was "important", "very important" or "extremely important" to the local economy a year ago to 93.78% indicating the same responses one year later.

Table 5 shows an overwhelming support to the Buy Local initiative even one year ago, which only increased over a year's time. 73.5% of respondents indicated that, one year ago, they were "supportive", "very supportive" or "extremely supportive" of buying local, which increased to 87.1% of respondents indicated the same responses to describe how supportive they are of buying local one year later.

Speaking of being supportive, additional survey questions reveal overwhelming support to local farmers and local small businesses. Tables 6 and 7 show 95.63% and 97.7% of respondents indicated being "supportive", "very supportive" or "extremely supportive" of local farmers and local small businesses, respectively.

TABLE 2: What do you think "Buy Local" means? (Check only one answer.)			
Buying products that are made in Guam from Chamorro-owned stores/businesses located in Guam.	176	41.12%	#1
Buying products INCLUDING both made on and outside of Guam from Chamorro-owned stores/businesses located in Guam.	58	13.55%	#3
Buying products that are made in Guam from stores/business located in Guam (including military base stores), even they are not Chamorro-owned.	76	17.76%	#2
Buying products INCLUDING both made on and outside of Guam from stores/business located in Guam (including military base stores), even they are not Chamorro-owned.	34	7.94%	
Buying products that are made in Guam from stores/business located in Guam but NOT on the military bases, even they are not Chamorro-owned.	26	6.07%	
Buying products INCLUDING both made on and outside of Guam from stores/business located in Guam but NOT on the military bases, even they are not Chamorro-owned.	28	6.54%	
Multiple answers	28	6.54%	
Other answers including *Buying from local farmers/buying local produce/vegetable*	2	0.47%	
responses	**428**		
skipped	**23**		

TABLE 3: A year ago, how aware were you of the idea of buying local? (Check one box only)
VERSUS Today, how aware are you of the idea of buying local? (Check one box only)

Awareness of Buy Local	A year ago		Today	
Not aware at all	71	16.47%	8	1.86%
Somewhat aware	143	33.18%	64	14.85%
Aware	107	24.83%	138	32.02%
Very aware	52	12.06%	105	24.36%
Extremely aware	58	13.46%	116	26.91%
AVERAGE (out of 5)		2.7285		3.5963
% AWARE, VERY AWARE, EXTREMELY AWARE		50.35%		83.29%
responses	**431**		**431**	
skipped	**20**		**20**	

TABLE 4: A year ago, how important did you think is buying local to our island? (Check one box only)
VERSUS Today, how important do you think is buying local to our island?

Importance of Buy Local	A year ago		Today	
Not important at all	17	3.95%	3	0.69%
Somewhat important	67	15.58%	24	5.53%
Important	127	29.53%	50	11.52%
Very important	63	14.65%	141	32.49%
Extremely important	156	36.28%	216	49.77%
AVERAGE (out of 5)		3.6372		4.2512
% IMPORTANT, VERY IMPORTANT, EXTREMELY IMPORTANT		80.47%		93.78%
responses	**430**		**434**	
skipped	**21**		**17**	

TABLE 5: A year ago, how supportive were you of buying local?
VERSUS Today, how supportive are you of buying local?

Supportive of Buy Local	A year ago		Today	
Not supportive at all	23	5.30%	3	0.69%
Somewhat supportive	92	21.20%	52	11.98%
Supportive	140	32.26%	116	26.73%
Very supportive	94	21.66%	134	30.88%
Extremely supportive	85	19.59%	128	29.49%
AVERAGE (out of 5)		3.2903		3.7581
% SUPPORTIVE, VERY SUPPORTIVE, EXTREMELY SUPPORTIVE		73.50%		87.10%
responses	**434**		**434**	
skipped	**17**		**17**	

TABLE 6: How important is it to support local farmers? (Check one box only)		
Not important at all	1	0.23%
Somewhat important	18	4.14%
Important	31	7.13%
Very important	122	28.05%
Extremely important	263	60.46%
AVERAGE (out of 5)		4.4437
% IMPORTANT, VERY IMPORTANT, EXTREMELY IMPORTANT		95.63%
Responses	**435**	
Skipped	**16**	

TABLE 7: How important is it to support your small businesses who operate in Guam? (Check one box only)		
Not important at all	0	0.00%
Somewhat important	10	2.30%
Important	42	9.66%
Very important	144	33.10%
Extremely important	239	54.94%
AVERAGE (out of 5)		4.4069
% IMPORTANT, VERY IMPORTANT, EXTREMELY IMPORTANT		97.70%
responses	**435**	
Skipped	**16**	

One measure of how effective an initiative has been is to find out to what extent it resulted in a desired outcome in the form of some action, a change in behavior. As mentioned earlier, in an initiative like the Buy Local initiative in Guam, the desired outcome is for local consumers to shift their purchases to local sources and away from non-local sources, which include the "Three O's": on-base (military stores), online and off-island. Our survey results suggest that the Buy Local initiative has been effective, with only 9.4% of the respondents indicating that the initiative has not encouraged them to purchase more from local vendors, leaving 90.6% who now purchase more from local vendors. Among the 90.6%, 22.65% indicated they "now purchase 41% to 60% more from local vendors", 18.31% "now purchase 61% to 80% more from local vendors" and 17.11% "now purchase 81% to 100% more from local vendors" instead of other sources.

Why support local buying?

As to the reasons why to support buying local, our survey shows the top three answers lead back to the economic benefits of a Buy Local initiative: multiplying local spending, creation of local jobs and generation of additional tax revenues. Allowing for multiple responses, these three reasons correspond to the following survey choices: "Locally owned businesses keep the dollars circulating on island" (23.08% of responses); "Hiring Guam residents first creates and maintains much-needed jobs" (18.03% of responses); and "Local businesses in Guam would directly contribute to Guam's tax base, adding much-needed revenue to support public service initiatives" (13.68% of responses).

TABLE 8: Compared to one year ago, by how much have you increased your purchases from local vendors instead of the military base, online source or shopping during your off-island trip(s)?

0% (no change in purchase decision)	39	9.40%	
I now purchase 1% to 20% more from local vendors	65	15.66%	
I now purchase 21% to 40% more from local vendors	70	16.87%	
I now purchase 41% to 60% more from local vendors	94	22.65%	#1
I now purchase 61% to 80% more from local vendors	76	18.31%	#2
I now purchase 81% to 100% more from local vendors	71	17.11%	#3
responses	**415**		
Skipped	**36**		

TABLE 9: What are good reasons for supporting buying local in Guam? (Check all that apply)

Locally owned businesses keep the dollars circulating on island.	23.08%	#1
Hiring Guam residents first creates and maintains much-needed jobs.	18.03%	#2
Shifting at least 10% of purchases from online, on-base, and/or off-island to local businesses will make a big difference in Guam's economy.	9.90%	
Small, local business owners in Guam are more likely to support Guam-based, non-profit organizations and events in the community.	11.79%	
Locally owned businesses in Guam make more local purchases, thereby requiring less transportation and contributing less to congestion and environmental pollution.	11.22%	
Local businesses in Guam are owned by people who live in the community, are less likely to leave and invest more in Guam's future.	12.30%	
Local businesses in Guam would directly contribute to Guam's tax base, adding much-needed revenue to support public service initiatives.	13.68%	#3
respondents	**1586**	

CONCLUSION

Our study shows that the Buy Local initiative in Guam has been effective one year since its marketing/educational campaign. Based on the survey of local residents, the initiative has been success in increasing local residents' awareness of buying local, their understanding of how important buying local is to Guam's island economy, and their support of the Buy Local initiative. Our study shows that the top reasons that local residents support buying local are to keep the dollars circulating/multiplying in the local economy, to create local jobs, and to increase local tax revenues.

More importantly, the initiative has been effective in changing local residents' spending behavior by shifting their purchases toward local businesses instead of buying from non-local sources such as the military base stores, online vendors or off-island vendors while shopping during a trip. The above evidence is expected to translate to stronger local demand for a wide array of goods and services and, consequently, greater employment opportunities in both the short-run and the long-run.

"Local Buying" is only one part of a three-part economic development strategy for Guam. The remaining two parts are "Local Hiring" and "Local Producing".

The economic benefits to the local community made possible by these greater employment opportunities could be maximized as more and more of these jobs are performed by local residents. It is true that U.S. labor laws clearly show a preference for "Local Hiring", i.e., employing qualified U.S. citizens and legal residents first before considering foreign workers for employment in the U.S. but, as noted earlier, small island economies are often faced with the challenge of limited productive resources, in this case, availability of skilled labor. Although in the short-run, some of the available jobs may go to foreign workers (whose spending toward the local economy is reduced by remittances sent to their home countries), long-run strategies (such as education and training) are in place to enhance the qualifications and availability of local residents to fill the available jobs.

The positive contribution of the Buy Local initiative to the local economy increases the more local businesses produce their own merchandise ("Local Producing") or the more they buy their merchandise from local sources instead of importing them. To this end, there have been efforts to increase local production of certain products that cater to local residents; for example, fruits and vegetables, including hydroponically grown lettuce and tomatoes grown by a local company called Grow Guam LLC as well as products for tourists and foreign markets (for example, Coco Jo's cookies and chocolates, which have been popular souvenirs for tourists but whose exports to Japan have grown since February 2012 (*Marianas Business Journal*, May 20, 2013)). More recently, the University of Guam-Pacific Center for Economic Initiatives has been exploring opportunities for "Local Producing" through the Guam Enterprise Initiative, which focuses on Guam's agriculture and identifies what products could be produced by each village along the lines of the One-Village-One-Product (OVOP) approach. According to Meyer as cited in Li and Schumann (2013), one way that the OVOP approach will benefit Guam's local economy by "increasing backward linkages to reduce leakages", which translates to increasing products that are supply locally instead of imported.

The economic development efforts discussed in this paper might be "business as usual" to policymakers and residents in larger, more sophisticated economies but to a small island economy like Guam and others, these efforts represent one of a limited number of options to try to move their economies forward. It might represent simply a "hope" for something positive or a "promise" of a better future but these hopes and promises could make a difference between a stagnant or, worse, a regressing economy and a slowly developing one.

ACKNOWLEDGEMENTS

Thanks to all inputs provided by the University of Guam-Pacific Center for Economic Initiatives team: Dr. Anita Borja Enriquez, Dr. Karri Trahin Perez, Dr. Annette Taijeron Santos, Dr. Fred R. Schumann, Ms. Eileen Agahan and Mrs. Natasha L. Cruz. Thanks to Ms. Eileen Agahan and Mrs. Natasha L. Cruz for facilitating the conduct of surveys at the Agana Shopping Center on November 10, 2012. Thanks also to Dr. Ron McNinch and Prof. Ramon Perez for encouraging the following students in their classes (listed in alphabetical order) to assist with conducting this survey: Luigi Bansil, Ferlyn Bulatao, Matthew Cabrera, Jennifer Cruz, Semjelfh Diva, Gabbie Franquez, Nicole

Garcia, Donita Giducos, Alpesh Godhania, Neelam Goetzfridt, Annie Joo, Ju Hee Kim, Kathleen Lamorena, Jhanice Miranda, Shanae Ooka, Joanna P., Kim Pacay, Michele Peele, Zchalyn Peralta, Anthony Quenga, Gilmore Rebujio, Renz Redolozo, Maria Romulo, Clarabelle Santos, R. Santos and Keo Torre. Thanks to Ms. Eileen Agahan and Mrs. Jaedene Tenorio for inputting the survey responses in Microsoft Excel worksheet.

REFERENCES

Buy Local Guam (2012, November). Learn Local event at Agana Shopping Center. Retrieved on December 22, 2013, from http://buylocalguam.org/learn-local-greeting-card-design-contest-winners

Buy Local Holiday Gift Guide (2011, December), a special publication by the *Pacific Daily News*.

Guam Bureau of Labor Statistics (2012). *Current Employment Statistics History, 1993-2012*

Li, Ning & Fred R. Schumann (2013). The One Village One Product (OVOP) Model and Economic Development on Guam. *Journal of Economics and Economic Education Research* 14(3), 21-34.

Marianas Business Journal (2013, May 20) by Paul Moroni. Coco Jo's cookie, chocolate exports on the rise.

Marianas Business Journal (2012, July 16) by Mar-Vic Cagurangan. Local community urged to explore new industries.

Marianas Variety (2012, December 26) by Louella Losinio. Historic Hagåtña corridor to be unveiled. Retrieved on December 11, 2013, from http://www.mvguam.com/local/news/27430-historic-hagatna-corridor-to-be-unveiled.html

Marianas Variety (2012, September 14) by Louella Losinio. Green Energy Career Expo. Retrieved on December 22, 2013, from http://mvguam.com/local/news/25843-green-energy-career-expo.html

Marianas Variety (2011, October 20) by Geraldine Castillo. University of Guam launches 'Buy Local' movement. Retrieved on December 22, 2013, from http://mvguam.com/local/news/20412-university-of-guam-launches-buy-local-movement.html

Pacific Daily News (2013, December 4) by Gaynor Dumat-ol Daleno. Celebrate historic barrio at block party in Hagåtña. Retrieved on December 22, 2013, from http://www.guampdn.com/article/20131204/NEWS01/312040014/Celebrate-historic-barrio-block-party-Hag-aring-t-ntilde-a

Pacific Daily News (2012, December 5) by Armando Cordoba. Buy Local plans cash mob Friday. Retrieved on December 22, 2013, from http://www.guampdn.com/article/20121205/NEWS01/212050313/Buy-Local-plans-cash-mob-Friday

Pacific News Center (2012, November 8). UOG PCEI to Host Learn Local Event at Agana Shopping Center. Retrieved on December 22, 2013, from http://www.pacificnewscenter.com/index.php?option=com_content&view=article&id=28870%3A-uog-pcei-to-host-learn-local-event-at-agana-shopping-center&catid=56%3Acommunity-events&Itemid=156

Pacific News Center (2012, September 26). UOG to Host "A Buy Local Affair" on FRIDAY. Retrieved on December 22, 2013, from http://www.pacificnewscenter.com/index.php?option=com_content&view=article&id=27571:uog-to-host-a-buy-local-affair-on-friday

Pacific News Center (2011, October 18). University of Guam Launches "Buy Local" Initiative WEDNESDAY. http://www.pacificnewscenter.com/index.php?option=com_content&view=article&id=17979%3Auniversity-of-guam-launches-qbuy-localq-initiative-wednesday&catid=56%3Acommunity-events&Itemid=156

Pacific News Center (2011, July 18). Guam Young Professionals to Hold Buy Local Mixer WEDNESDAY. http://mail.pacificnewscenter.com/index.php?option=com_content&view=article&id=15647%3Aguam-young-professionals-to-hold-buy-local-mixer-wednesday&catid=56%3Acommunity-events&Itemid=156

Ruane, Maria Claret M. (2011). Macroeconomic Multipliers: The Case of Guam. *Journal of Pacific Studies* 31(1), 141-155.

Ruane, Maria Claret M. (2011, December). Estimating the Spending Multiplier in Guam. *PCEI Technical Report Issue 15*. University of Guam-Pacific Center for Economic Initiatives. Available at http://www.uogpcei.org/docs/guam/pcei-technical-reports/PCEI-TechnicalReport-Issue15-GuamSpendingMultiplier-December2011.pdf

Ruane, Maria Claret M. (2012, April 13). "Buying Local: A Way to Increase Guam's Spending Multiplier", Presented at the Pacific Hotel and Restaurant Expo (PHARE) in Tumon, Guam.

U.S. Department of Commerce, Bureau of Economic Analysis (2012, September 24). Press Release: The Bureau of Economic Analysis (BEA) Releases Estimates of Gross Domestic Product, Gross Domestic Product by Industry, Compensation by Industry and Detailed Consumer Spending for Guam. http://www.bea.gov

APPENDIX A: EXAMPLES OF EVENTS AND ACTIVITIES AIMED AT PROMOTING THE BUY LOCAL INITIATIVE

- The Guam Chamber of Commerce Small Business Committee hosted a seminar in conjunction with the Small Business Expo on June 30, 2011. The theme of the expo and the seminar is "Buy Local."
- The Guam Young Professionals, a committee of the Guam Chamber of Commerce, held a "Buy Local" mixer on July 20, 2011 (*Pacific News Center*, July 18, 2011)
- An announcement on September 1, 2011 by the U.S. Department of the Interior-Office of Insular Affairs for the $150,000 initial grant to the UOG-PCEI and GEDA, paving the way to the launch of the "Think, Support and Buy Local" is the marketing campaign of the Buy Local initiative in Guam that was officially launched on October 19, 2011 at the Bernardo's Dragonfruit Farm in Chalan Pago. This campaign is an economic development strategy to promote awareness of the benefits of supporting local (Guam-based) businesses and hiring. . Dr. Enriquez said that "the 'buy local' movement is not a new one" but just "… hasn't built momentum in Guam until now." (*Pacific News Center*, October 18, 2011)
- *Buy Local Holiday Gift Guide* was published by the *Pacific Daily News* in December 2011. According to Mr. David Leddy, President of the Guam Chamber of Commerce, "By using this resource Guide and discovering the wonderful products and services available locally, you help support our economy. Dollars spent locally helps to create and protect jobs, keep businesses open, provide needed funds for government services and enhances economic development."
- "Listen Local: An iHeart Local Music Event" to showcase local music talents on January 28, 2012
- "Buy Local Commitment", a company's pledge to "Think, Support and Buy Local" on February 29, 2012. As of December 21, 2012, sixty-eight (68) local companies, government agencies and local organizations made the pledge to support and sign the Buy Local commitment on a monthly basis. See Appendix A for a list of these Buy Local organizations.
- "Maila' Ta Fan Chesa" on March 23, 2012. Hosted by the Micronesian Chefs Association, the event featured local chefs who created appetizers, desserts and beverages using fresh fruits, vegetables and herbs from local village farmers.
- "Why Local Exhibit and Forum" at the Pacific Hotel and Restaurant Expo on April 12-13, 2012. (http://www.buylocalguam.org)
- "Savor Guam Food Festival", which celebrated sustainability and the use of local products and resources on April 22, 2012

- "Think, Support, Buy Local at Chamorro Village" campaign and contest from June 22 to July 18, 2012 aimed at promoting thinking and buying local by supporting the Guam Product Seal program participants, Chamorro Village tenants and vendors, and BuyLocalGuam businesses. (http://www.buylocalguam.org). Note that "Chamorro" refers to the native people of the island of Guam and its culture.
- "A Buy Local Affair" on September 28, 2012 at the UOG-School of Business and Public Administration Building. The event showcased products made in Guam such as artwork, music, packaged edible goods and more from Guam Product Seal businesses, Buy Local vendors, Isla Center for the Arts, and other businesses and organizations and was part of the University of Guam's 60th Year Anniversary celebration. (*Pacific News Center*, September 26, 2012)
- "Learn Local: Totes for Teachers" in October 2012. Guam Department of Education fifth-grade teachers and students received tote bags filled with items, such as folders, magnets and stickers, to help promote the initiative. The objective is to reach households in order to help families understand the value of supporting Guam-based businesses and the resulting positive multiplier impact to the local economy. (http://www.buylocalguam.org)
- "Greeting card design contest" in October 2012 on the theme of "What does Christmas/Holiday in Guam mean to me?" The winning designs had their artwork featured on Zories Only greeting cards and sold as a local product. Zories Only will then donate a portion of greeting card sales to the schools, thereby supporting public education. (http://www.buylocalguam.org)
- "Cash Mob" on December 7, 2012 where "a group of people armed with at least $5 spend that money at a designated local business". (*Pacific Daily News*, December 5, 2012)
- Unveiling of "San Ignacio Barrio," the first historic Hagåtña corridor, by the University of Guam's Pacific Center for Economic Initiatives, in collaboration with the Guam Preservation Trust, on December 27, 2012 (*Marianas Variety*, December 26, 2012)
- Launch and Block Party at the San Ignacio Barrio, in collaboration with the The University of Guam Pacific Center for Economic Initiatives, in collaboration with the Guam Preservation Trust, Hagåtña Restoration and Redevelopment Authority of the Department of Chamorro Affairs, Guam Humanities Council and Local First! Guam, on December 5, 2013 (*Pacific Daily News*, December 4, 2013)

APPENDIX B: LIST OF LOCAL COMPANIES, GOVERNMENT AGENCIES AND ORGANIZATIONS THAT HAVE MADE THE BUY LOCAL PLEDGE AS OF DECEMBER 10, 2013

1. A&L Crafts
2. Agana Shopping Center
3. Amot Taotao Tano Farm
4. ASC Trust Corporation
5. Black Tie Events & Floral Designs
6. Bensin Guam Enterprises, Inc.
7. Budget Car Rental
8. Cars Plus
9. Chamorro Island Bar B.Q.
10. Cham's Cuisine
11. Che'lu
12. Cycles Plus
13. Construction Resources
14. Cutie Petunias
15. Deloitte & Touche, LLP
16. Dipstixx
17. Docomo Pacific
18. Eleventh Essence
19. First Hawaiian Bank
20. Freedom Air
21. Galaide Group, LLC
22. GPSI Guam
23. Grow Guam
24. Guam Cornerstone
25. Guam Chocolate and Pastry
26. Guam Reef Hotel
27. Guam Style
28. Guam Strategic Development, LLC
29. Hava Java Cafe
30. Horizon Properties, Inc.
31. Ifit Addao yan Famaguon
32. Ifil Shop
33. IMG Studio, LLC
34. Inspire Ad Agency
35. International Distributors, Inc.
36. Island Pacific (I.P.) Coffee Shop
37. J&F Tropical Hot Rod Hut
38. Kaduku Cards and Shirts
39. Ken & Dan & Flo Farmers Market
40. Kristal Kollection

41. Loco Promos
42. M-80 Systems, Inc.
43. MainStreet Delicatessen & Bakery
44. Market Wholesale Distributors, Inc.
45. The Mermaid Tavern & Grille
46. Moda Gino's
47. My Secret Garden
48. New Memories
49. O&M Safety Analysis
50. Ohana Hawaiian BBQ
51. Onward Beach Resort
52. Outback Steakhouse
53. Pacific Human Resource Services, Inc.
54. Paradise Auto Spa
55. Payless Car Rental
56. Perez Bros., Inc.
57. Pika's Café
58. Puppy Love Guam
59. Salon Paradis
60. Security Title, Inc.
61. SEI Guam
62. Sky Dive Guam
63. The Occasion
64. Thrifty Car Rental
65. Tony Roma's Restaurant
66. Triple J Enterprises
67. Triple J Five Star Wholesale Foods, Inc.
68. United Airlines
69. Xerox Corporation
70. Guam Chamber of Commerce
71. Guam Economic Development Authority (GEDA)
72. UOG Pacific Center for Economic Initiatives (PCEI)
73. Guam Contractors Association
74. Guam Hotel and Restaurant Association

APPENDIX C: SURVEY RESPONDENTS' PROFILE

Gender	% of total respondents
Female	67.92%
Male	32.08%
Respondents	424

Age	% of total respondents
18-25	19.81%
25-30	12.89%
31-40	20.05%
41-50	22.20%
51-60	16.23%
61-65	4.30%
65+	4.53%
Respondents	**419**

Ethnicity	% of total respondents
Chamorro	54.69%
Other Pacific	7.28%
Filipino	19.95%
Other Asian	3.05%
Caucasian	5.63%
African-American	0.23%
Others	3.05%
Mixed ethnicity	6.10%
Respondents	**426**

Marital Status	% of total respondents
Single	44.34%
Married (including common-law unions)	50.94%
Others	4.72%
Respondents	424

Annual Household Income (in $1000s)	% of total respondents
0-25	36.88%
25-50	29.38%
50-75	14.38%
75-100	8.75%
100+	10.63%
Respondents	**320**
Average household income	**$51,480**
Median household income	**$36,000**

Number of Dependents in the household	% of total respondents
0	28.02%
1	18.13%
2	15.38%
3	14.56%
4	9.89%
5	5.49%
6	4.95%
7 or more	3.57%
Respondents	**364**

Occupation/Line of Work	% of total respondents
Banking, Insurance, Financial Services	6.55%
Communications/Utilities	2.02%
Construction	2.77%
Education	4.28%
Federal Government	8.82%
GovGuam	19.90%
Hotels	2.27%
Medical/Health Services	6.05%
Military	3.78%
Restaurant	6.05%
Retail Trade	8.82%
Services	7.30%
Tourism/Tour Agencies	1.26%
Wholesale Trade	1.76%
Self-Employed	1.76%
Non-Profit/Community Organizations	2.02%
Retired	5.04%
Students	3.23%
Unemployed/Homemaker	4.79%
Others	1.51%
Respondents	**397**

Education Level	% total respondents
Some high school	5.50%
High school graduate	33.49%
Some college	25.36%
College degree	25.60%
Master's degree or higher	10.05%
Respondents	**418**

Years living in Guam	% total respondents
less than 1 year	3.52%
1 to 3 years	2.58%
3 to 5 years	2.82%
5 to 10 years	6.34%
More than 10 years	84.74%
Respondents	**426**

THE MARKET STRUCTURE DYNAMICS CREATED BY DE-BUNDLING OF AIRLINE BAG FEES

Hans O. Schumann, Texas A&M University – Kingsville
Harmeet Singh, Texas A&M University – Kingsville

ABSTRACT

Charging fees for checked bags is a recent phenomenon in the airline industry. Until late 2008, checking bags was a free service, the cost of which was bundled in base airfare. Our analysis of airline markets before and after the implementation of bag fees validates Chen's (1997) model of mixed bundling strategies leading to market differentiation and weakening of Bertrand equilibrium.

Keywords: Pricing, Market Structure Dynamics, Oligopoly Pricing, Baggage Fee, Revenue Management, Industrial Organization.

JEL Code: L11, L13

INTRODUCTION

Charging Baggage fees for checked bags is a recent phenomenon in the airline industry. Until late 2008, checking bags was a free service, the cost of which was bundled in the base airfare. This article builds on the findings by Henrickson and Scott (2012) and Schumann and Singh (2013) that showed airlines continuing the 'bags fly free' tradition were more successful in increasing base airfares than competitors charging checked bag fees. Henrickson and Scott (2012) and Schumann and Singh (2013) both base their findings on U.S. Department of Transportation (USDOT) 10 percent sample collected from all domestic airline tickets sold in United States, the former having looked at a sample of top grossing markets while the latter focused on sample of markets in which both "Fee" (airline carriers that charged fee for any and all checked baggage) and "No-Fee" (airline carriers that allowed at least one checked baggage at no additional charge with the purchase of base fare) carriers were jointly present. This article extends the research to specifically evaluate markets in which all carriers opted to de-bundle the air fare by charging separately for checked baggage. Looking at the two types of markets we noticed significant differences in the airline competitive dynamics.

The decision for some airlines to start charging for baggage is especially intriguing given that airlines have historically been very good at developing fences to separate leisure and business travelers, the former of which are assumed to be significantly more price elastic than business travelers as "many price-inelastic consumers travel for business related reasons and tend to value their time highly" (Gerardi and Shapiro 2009). Baggage fees would appear to have greater impact

and raise effective prices for leisure travelers, the opposite of what airlines would be expected to do.

The results of this study could extend well beyond the airline industry, as there are pricing structure parallels to many other industries where oligopolistic or duopolistic firms have an ability to bundle their products into a combined price, or to price and sell their products or product features "a la cart".

Different firms in all types of industries follow multiple pricing and bundling strategies. One extreme is the "a la cart" structure in which each individual attribute of the product is priced individually (de-bundled), while the other is where a product is offered 'all inclusive' with one price. "All inclusive" pricing structure sometimes have an adverse consumer reaction because some attributes of the product may have very little or no value to the consumer. Alternatively, other consumers may become upset by being asked to pay separate prices for each product attribute and complain about being "nickel and dimed".

Most economic analysis of bundling has been done in a monopolistic setting. Understanding optimal bundling strategies in an oligopolistic setting is more difficult as competitors could under-cut any bundled offering, yielding only the Bertrand equilibrium on the "bundle". Additionally, the existence of competitors can keep a potential product bundler from using mixed bundling strategy that relies on artificially inflated single product prices. We analyze the air fare bundling/de-bundling strategy in the US domestic airline industry which is oligopolistic in its structural attributes.

REVIEW OF LITERATURE

Henrickson and Scott (2012) and Schumann and Singh (2013) presented analysis of USDOT airline ticket data to analyze the de-bundling of ancillary services in the US domestic airline industry, in particular the baggage fee. Before that the popular and financial press has been the only sources investigating this phenomenon.

Burstein (1960) and Adams and Yellen (1976), were among the first to study product bundling in a monopoly setting. Carbajo, Meza, & Seidmann (1990) and Chen (1997) identified that bundling is often seen in duopoly or oligopolistic context. Specifically, Carbajo et. al. (1990) cites the photo film and processing industry as an example in which both the film and processing are sometimes bundled and at other times sold separately as individual offerings. Chen's (1997) examples include computer hardware and software, credit and travel companies with loyalty programs. Schumann (1986) explained that airline frequent flier programs are essentially a way in which oligopolistic airline carriers bundle their products together for greater differentiation.

Phillips and Schutte (1988) wrote about the bundling strategies adopted in gasoline station markets. One can draw parallels between the gasoline station markets and the airline industry. Full service gasoline stations bundled "full service" with the price of gasoline. This is similar to airlines providing baggage service as part of the base air fare purchase. Some gasoline stations choose a mixed pricing strategy wherein they offer some self-service pumps and some pumps with "full service" at a higher price. This is similar to the airlines offering an airfare that does not include checked baggage service and the baggage service is offered with a separate price. In the gasoline station markets there is another strategy adopted more commonly by most gasoline

stations these days, is completely self-service pumping of gasoline. With the exception of Eastern Airlines "moonlight special" flight service in the 1980s, that the airline did not guarantee checked baggage availability, the airline equivalent of a completely self-service gasoline station as of yet does not exist[1].

Phillips and Schutte's (1988) focus in researching the gasoline station markets was on determining the cross-price elasticities of full and self-serviced customers. Our focus in this article is to determine how a similar price partitioning affected the competitive structure and overall price levels within the airline oligopoly.

Chen (1997) was one of the first to develop a theoretical model that finds one duopolistic firm choosing a bundled product pricing strategy where the other offers an 'a la cart pricing'. The current market example of airline industry with their differing baggage fees mirrors this structure.

Product bundling and de-bundling is not only of interest to economists, but is also studied in the marketing field. Price Partitioning (Schindler 2011) is the term used in the marketing literature to describe when a firm de-bundles the product attributes to offer them via "a la cart" pricing. Customers not demanding a product attribute are not required to pay for that, when the price is "partitioned" among multiple features of the product. Price partitioning can also be used to communicate to the consumer the added cost of providing multiple product attributes. For example, an Ultra-Low Cost Carrier (ULCC) in United States separates the total fare in multiple cost factors the passenger is being asked to pay for including, the cost of fuel and "government's cut"[2].

Providing the air passenger with the service of checked baggage has costs of its own. These costs include, but are not limited to, handling, cargo space, security screening, etc. Similarly, there are costs to allow carry-on bags to board with the passenger. The most apparent being the time delay caused during boarding and disembarking of planes which decreases the scheduled asset productivity of the aircraft. Currently, the ULCCs operating in the United States domestic airline routes, have implemented in their pricing structure a higher carry-on baggage fee than their checked-baggage fee[3]. This clearly indicates that they wish to encourage passengers to check bags rather than bringing those as carry-on implying that, in their opinion, the cost of carry-on may be higher than checking a bag.

AIRLINES INDUSTRY AND THE ECONOMIC CONDITIONS DURING THE STUDY PERIOD

During the period of study there were major changes in the economy which affected the market for air travel. The great recession of 2007-09 had a negative impact on overall economic income levels and for services such as air travel. Indeed, precisely because of this reduction that some carriers chose to consider charging for bags, at least this was the public claim by some airline executives (AMR Corp., 2009, Delta Air., 2009 and UAL Corp., 2009). Because of these changes in the macro-economic environment, it is very difficult to analyze any changes in demand caused specifically by price changes. However, since all carriers operated in the same economic environment, we can compare airline to airline in terms of their pricing strategy and the number of passengers carried.

As we see in Figures I and II, 2008 brought significant economic changes that impacted the airline industry. The Producer Price Index (PPI) for Jet Fuel which constitutes a major variable cost for keeping the airplanes in the air was on the rise during all the "Prior" period (Defined as study period ranging from the beginning of first quarter of 2006 until the end of second quarter of 2008). During the same time, the overall PPI for Scheduled Passenger Air Transportation rose as well. These cost increases, along with the contracting economy, were significant in influencing most airlines to implement a new stream of ancillary revenue, in this case initiation of checked baggage fee. With the combination of the weakening economy together with the explosive growth in fuel expenses, airlines were looking for ways to protect their income. Indeed, it was these factors which were publically cited by the airlines in their announcements of initiating baggage fees.

PRODUCT DIFFERENTIATION IN AIRLINE INDUSTRY

The Airline industry is a very interesting industry to study pricing effects for a variety of reasons. Foremost, the basic transportation services provided by the carriers are identical with relatively little differentiation with each airline providing the same service of moving a customer from location A to location B. However, it is not a pure commodity as some minor differences do exist such as, flight schedule, seat comfort, quality of food snacks provided (if any), and any participation in frequent flyer loyalty programs.

The airline industry serves a heterogeneous population made up of vacation and leisure travelers. Since de-regulation, airlines have become very adept at developing pricing tools to segment price sensitive leisure travelers from price insensitive business travelers. These include fences such as, Saturday night stays, non-refundable fares, in-advance purchase requirements, etc., which have historically allowed airlines to offer leisure travelers a low fare while simultaneously charging the price inelastic business travelers a much higher price for the same base transportation service. As mentioned above, it is within this context that the current bag fees decisions are of such interest. As one would imagine, such a policy (fare increase) would be disproportionately aimed at a leisure traveler rather than a briefcase-carrying business traveler. The issue whether this policy disproportionately affects one segment greater than the other is left as an issue for further research.

Because the US airline industry used to be federally regulated, the old Civil Aeronautics Board collected extensive data off of a 10 percent sample of every airline ticket sold in United States. Even after de-regulation, the Department of Transportation has continued this data collection tradition. Very few industries are so data rich when it comes to market and price information.

THE DATA

The data for this analysis comes from the US Department of Transportation. The DOT requires airlines to submit a 10% sample of all domestic airline tickets sold with all data related to itinerary and fare charged. This data is compiled quarterly and made available via the internet to researchers and the public. For each origin-destination city pair and major carrier, the published

DOT data fields include: Average Fare, Number of Passengers, and carrier Market Share. For the airline's data to be included, they must have a market share of over 10% during that quarter. An origin-destination city pair includes all one-way segments in either direction regardless of their routing. For instance, data from a ticket for a flight from Chicago to Cleveland connecting to a flight from Cleveland to Boston would be included in the Boston-Chicago city pair and would not appear in either Chicago-Cleveland or Cleveland-Boston city pairs.

From this data, we analyze distinct city pairs for which there was consistent competition during the entire 2006 to 2010 study period. Markets for which the same set of competitors remained during the entire study period were retained, while those with only one dominant carrier or those with significant changes in competitors serving the market were dropped with the exception of certain instances of airline name changes due to airline mergers. Our final data set consisted of 112 distinct Origin/Destination city pairs.

Two major airline mergers happened during the study period, Delta/Northwest and US Airways/America West. Where possible we re-coded and combined data from both airlines for the period prior to their merging of operations. In some instances this may cause the market share gains of these two carriers to be over-reported. This would happen if one of the two pre-merger airlines had unreported data in a market due to falling under the 10% market share threshold.

Our analysis focuses on changes between two of the three time periods in our study, the "Prior" and "After" periods. Respectively these are defined as 1Q06 – 2Q08, and 1Q09 – 4Q10. Data for the "Transition" period, 3Q08 – 4Q08, is ignored as this was the time when most carriers were implementing baggage fee starting first checked baggage.

The carriers were broken into two categories.

"Fee" carriers refer to the US airlines that began charging separately for checking a first bag in 2008.

"No Fee" carriers are those carriers that continue to allow at least one "Free" checked bag during the entire study period.

Table I identifies the carriers, their classification, the bag fee charged and the date of this policy change. It also depicts the estimated bag fee revenue collected per passenger.

As described above, each of the city pair markets has had at least two competitors competing for passengers during the study period. We have broken these markets into two categories.

"Fee/Fee" markets are those where the major competitors all fit into the "Fee" carrier category. There are 46 markets that fall into this classification.

"Fee/Free" markets are those where there are competitors from both the "Fee" and "No Fee" carrier categories. There are 66 markets that fall into this classification.

Theoretically, there also could exist a third category, "Free/Free". Unfortunately, we could include none of the markets in which two or more "No Fee" carriers competed because they did not meet the criteria that competition must have existed during the "Prior" periods as well. Additionally, we did not look at markets that were dominated by one carrier.

RESULTS

When we look at the subset of markets where both "Fee and "No Fee" carriers operated we saw that both types of carriers were able to increase their base airfares despite the headwind from the economic environment. This price increase averaged $13.39 and was $16.69 and $11.31 for the "No Fee" and "Fee Carriers", respectively (see Table II). Although the "Fee" carriers had a lower price increase, they had access to the bag-fee revenue stream that the "No Fee" carriers did not.

When we analyze the markets where only "Fee" carriers operated, the results are remarkably different. Instead of fares increasing, the average base fare decreased by $12.28 during the exact same time periods. As the analysis is using the same data during the time period with the same macro-economic factors pertaining to the U.S. domestic airline industry, the changes in base fare is a valid comparative metric between the two types of markets. The overall difference in fare change for "Fee" carriers between these two types of markets is a little over $23.59. Table II details these findings for all major airline carriers.

As there is a positive change in the markets where "Fee" and "No Fee" carriers competed, and all factors are identical, our only conclusion is that there must be a difference between the competitive dynamics in the two market types.

This situation may be exactly what Chen (1997) predicted. His theoretical model predicted that if one oligopolistic firm may choose to bundle while the other did not, this would result in a differentiated market where price levels increase as the Bertrand competition was significantly weakened. In this case, baggage fees appear to be used as this creator of product/service differentiation leading to pricing power for all participating firms in the market.

Airlines do not disclose the baggage per route and the DOT only collects base airfare data. We therefore do not know how much more ancillary baggage revenue, if any, is collected when a "Fee" carrier does not have a "No Fee" alternative. Table I shows that even the largest system-wide revenue per passenger is dwarfed by the markets' base-fare difference. Therefore, even if there were differences in customers' propensities to check bags in these two markets, the resulting extra revenue increase could not close the base fare difference we have seen.

The result is clear, airlines that have charged bag fees benefit when their competitors choose not to. From Schumann and Singh (2013), airlines that choose not to charge bag fees when their competitors do, appear to post economic gains equal to or better than their "Fee" counterparts.

CONCLUSIONS

A priori, we had thought that the existence of carriers not charging for bag fees might be detrimental to the carriers who do. The results however indicate that "Fee" carriers do much better competing with a "No Fee" carrier vs. competing with another "Fee" carrier. This however does not appear to be at the expense of the "No Fee" carrier as both showed an ability to increase fares.

While we originally began this analysis to identify which airline executive group made the "right" bag-fee decision, it appears that collectively they both did. Should the "No Fee" carriers

join their counterparts, (or vice versa) this analysis predicts that both carrier types would stand to lose pricing power.

REFERENCES

Adams, W. & J. Yellen (1976). Commodity Bundling and the Burden of Monopoly." *The Quarterly Journal of Economics, 90*(3), 475 – 498.

AMR Corporation, 2009 SEC Filings: Annual Report: 10-K, Dec 31, 2008, p. 30, from AMR Corporation investor relations website. Retrieved on June 18, 2013, from http://phx.corporate-ir.net/phoenix.zhtml?c=117098&p=irol-sec&secCat01.1_rs=11&secCat01.1_rc=10&control_searchbox=&control_selectgroup=1

Burstein, M. L. (1960). The Economics of Tie-in Sales. *Review of Economics and Statistics, 42*(1), 68–73.

Carbajo, J., D.D. Meza & D.J. Seidmann (1990). A Strategic Motivation for Commodity Bundling. *The Journal of Industrial Economics, 38*(3), 283-298.

Chen, Y. (1997). Equilibrium Product Bundling. *The Journal Of Business, 70*(1), 85-103.

Delta Air Lines, Inc. 2009 SEC Filings: Annual Report: 10-K, Dec 31, 2008, p. 34, from Delta Air Lines, Inc. investor relations website. Retrieved on June 18, 2013, from http://phx.corporate-ir.net/phoenix.zhtml?c=71481&p=irol-sec&control_selectgroup=Annual%20Filings

Gerardi, K.S. & A.H. Shapiro (2009). Does Competition Reduce Price Dispersion? New Evidence from the Airline Industry. *Journal of Political Economy, 117*(*1*), 1-37.

Henrickson, K.E. & J. Scott (2012). Baggage Fees and Changes In Airline Ticket Prices. *Pricing Behavior and Non-Price Characteristics in the Airline Industry Advances in Airline Economics, 3*, 177-192.

Phillips, O. R. & D.P. Schutte (1988). Identifying Profitable Self-Service Markets: A Test in Gasoline Retailing. *Applied Economics, 20*(2), 263-272.

Schindler, R.M. (2011). *Pricing Strategies: A Marketing Approach*, Thousand Oaks, CA: Sage Publications, Inc.

Schumann, H. O. (1986). *Oligopolistic Nonlinear Pricing – Trading Stamps and Airline Frequent Flyer Programs*. Unpublished doctoral dissertation, Northwestern University.

Schumann, H. O. & H. Singh (2013). The De-bundling of U.S. Airline Bag Fees and Fares: An Exploratory Study. *Journal Of The Academy Of Business & Economics, 13*(1), 91-98.

UAL Corporation and United Air Lines, Inc., 2009 SEC Filings: Annual Reports: 10-K, Dec 31, 2008, p. 39, from United Continental Holdings, Inc. investor relations website. Retrieved on June 18, 2013, from http://ir.unitedcontinentalholdings.com/phoenix.zhtml?c=83680&p=irol-sec&secCat01.1_rs=31&secCat01.1_rc=10&control_searchbox=&control_selectgroup=1

APPENDIX A

Figure I

Figure I provides a visual comparison between "Prior" and "After" study periods in terms of macroeconomic variables in United States.

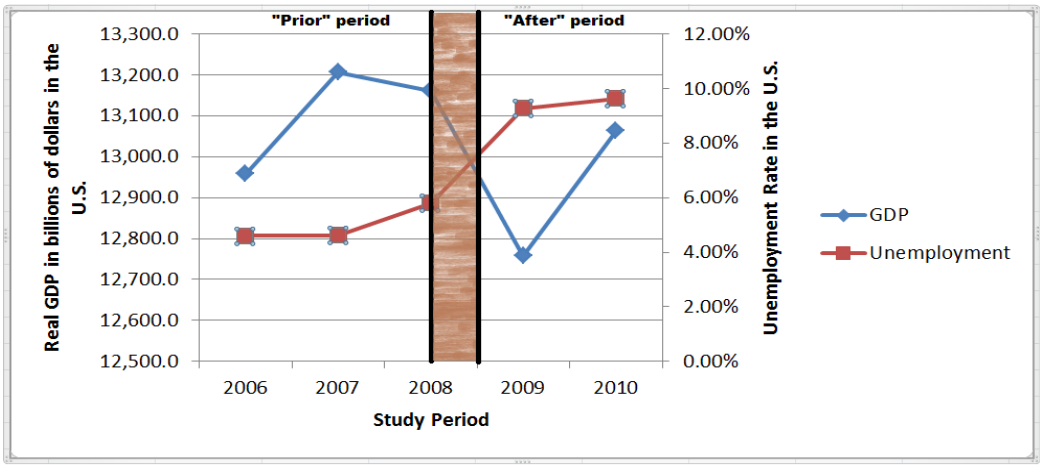


Figure II

Figure II provides a visual comparison between "Prior" and "After" study periods in terms of airline industry Producer Price Indexes.

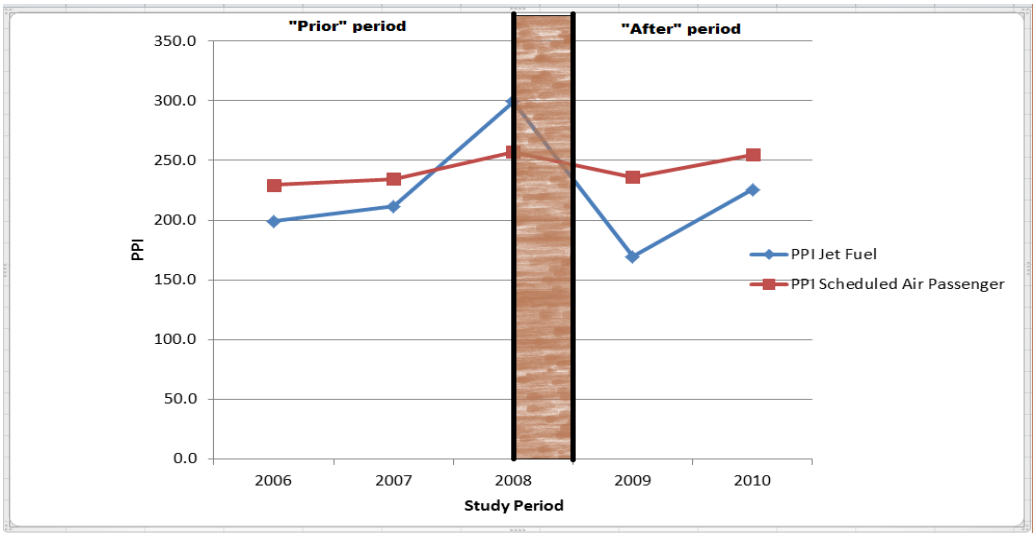


Table I

Table I shows the time line when each of the major "Fee" carriers initiated extra fees for first and second checked baggage. Prior to these dates, all of these carriers allowed for at least 1 free checked bag.

Airlines	Checked Bag Fee: 1st and 2nd	Bag fee Initiated	*Estimated Bag fee revenue/passenger[4]
Delta (DL)/ Northwest (NW)	$23, $32	2008 Q4	$8.58
American (AA)	$25, $35	2008 Q2	$6.74
United (UA)	$23, $32	2008 Q3	$5.79
US Airways (US)/ America West (HP)	$23, $32	2008 Q3	$9.91
Continental (CO)	$23, $32	2008 Q4	$7.86
AirTran (FL)	$15, $25	2008 Q4	$6.20
Midwest (YX)	$20, $30	2008 Q4	Not Reported
Frontier (F9)	$20, $30	2008 Q4	Not Reported
Average			$7.51

*It must be noted that these baggage fees are for the carrier as a whole and are not available on a route-by-route bases. As economic theory would dictate, the percentage of passengers checking bags on "Fee" carriers may be lower in the "Fee/Free" markets where customers expecting to check bags could book away from the carrier.

Table II

Table II provides detailed analysis of base fare comparison between the "Prior" and "After" study periods. Fare change between the two types of markets: "Fee/Fee" and "Fee/Free" is presented in the table as well.

	**Average of Airline Fare "Prior"	**Average of Airline Fare "After"	Average of Change in Fare "After – Prior"	Average of Percentage Change in fare	Count of Markets	**Average of Quarterly Airline Average Fare
Fee Markets	**174.81**	**162.53**	**-12.28**	***-7.73%**	**46**	**171.58**
Fee Carriers	174.81	162.53	-12.28	*-7.73%	46	171.58
AA	198.16	196.93	-1.23	-1.41%	13	200.62
CO	205.67	209.07	3.40	1.86%	2	210.08
DL***	185.10	163.02	-22.08	-11.98%	28	176.83
FL	134.16	121.79	-12.368	-9.30%	1	130.38
UA	227.51	239.18	11.67	5.24%	2	238.00
US***	209.83	197.83	-11.99	-6.57%	0	207.43
YX	134.82	122.80	-12.02	-8.92%	0	133.92
Fee/Free Markets	**155.03**	**168.42**	**13.39**	***9.75%**	**66**	**161.88**
Fee Carriers	164.60	175.91	11.31	*7.67%	60	170.37
AA	158.55	170.05	11.50	7.96%	33	164.32
CO	185.86	199.14	13.28	8.73%	6	192.70
DL***	168.42	171.62	3.20	1.19%	3	170.29
F9	120.76	114.94	-5.81	-4.57%	2	116.44
FL	118.36	119.93	1.57	1.36%	4	118.92
UA	176.59	187.71	11.12	7.37%	8	182.89
US***	163.09	186.19	23.10	16.35%	4	173.47

Table II
Table II provides detailed analysis of base fare comparison between the "Prior" and "After" study periods. Fare change between the two types of markets: "Fee/Fee" and "Fee/Free" is presented in the table as well.

	**Average of Airline Fare "Prior"	**Average of Airline Fare "After"	Average of Change in Fare "After – Prior"	Average of Percentage Change in fare	Count of Markets	**Average of Quarterly Airline Average Fare
No Fee Carriers	139.82	156.51	16.69	*13.04%	6	148.38
B6	143.64	143.50	-0.14	-0.79%	6	144.84
U5	142.92	138.35	-4.57	-3.20%	0	141.43
WN	139.29	158.46	19.17	15.05%	0	148.95
All Markets All Carriers	**162.60**	**166.17**	**3.57**	***3.06%**	**112**	**165.59**

* Statistically significant difference at α=0.05 level
** In Dollars.
***DL includes NW and US includes HP.

ENDNOTES

Schmitt, Eric, "Overnight Flight: Bargain for Spontaneous Travelers." The New York Times, Mar 09, 1987, Archives, URL: http://www.nytimes.com/1987/03/09/nyregion/overnight-flight-bargain-for-spontaneous-travelers.html accessed on June 03, 2013.

As of June 14, 2013, the Total Price breakdown provided for a flight from Los Angeles to Houston on web page of Spirit Airlines Inc. URL: http://www.spirit.com/DPPCalendarMarket.aspx.

As of June 14, 2013, the baggage fees information provided on web page of Spirit Airlines Inc. URL: http://www.spirit.com/OptionalFees.aspx. and on web page of Allegiant Travel Company. URL: http://www2.allegiantair.com/popup/optional-services-fees#baggage.

Yamanouchi, Kelly, "Delta No. 2 in bag fees per passenger." *The Atlanta Journal-Constitution,* Jun 15, 2011, Retrieved August 3, 2012, from URL: www.ajc.com/news/business/delta-no-2-in-977976.html accessed on August 3, 2012.

SCAREDY CATS TO COOL CATS: HOW TIME PERSPECTIVE MATTERS IN ATTITUDE AND INTENT TOWARD FINANCIAL DECISIONS

Marilyn Spencer, Texas A&M University-Corpus Christi
Valrie Chambers, Texas A&M University-Corpus Christi
Bilaye Benibo, Texas A&M University-Corpus Christi

ABSTRACT

The widely broadcast evidence of a looming economic downturn in late 2008, with a magnitude and length most of us had never before experienced, provided an unusual opportunity to measure the financial decisions of middle income, relatively well educated Americans in a time of financial crisis. The degree of devastation and even panic some "scaredy cat" individuals felt about their financial security and their financial future during the economic meltdown in the late summer and fall of 2008, and the continually higher unemployment rates until late 2010, surely affected decisions about saving and investing funds. This study utilizes the Zimbardo Time Perspective Index (ZTPI) developed by Zimbardo and Boyd (2008), to measure the influence of time perception on financial decisions in that very uncertain environment, while statistically correcting for demographic influences. Although numerous studies have attempted to explain what propels an individual's decisions concerning how much to spend or save, and how risk seeking or risk averse they are, none have utilized the psychology of time perspective. The two original issues examined in this study are (1) the intent to change jobs, and (2) the intent to move funds. As the recession lingered well beyond fall 2008, a refinement of the instrument was used to examine intentions concerning job changes. (Funds would likely already have been moved.) Using the ZTPI questions as a starting point, this study created and tested a second set of tailored questions that included 21 questions specific to time perspectives of financial issues, in an attempt to provide a more accurate picture of the influences of time perspective on the intent to change jobs. Then in late 2012, a third set of data was tested, using 60 modified-ZTPI items. By that time the economy had stabilized and unemployment was inching downward. Those results showed less predictability than when the job market was in crisis, indicating the return to a less emotional, "cool cat" decision-making process. However, taken together, these three sets of results indicate both the promise of time perspective on the intent to change jobs, and the usefulness of questions that more directly measure time perspective with regard to finances, in times of general financial crisis.

Key words: time perspective, job, asset allocation, risk

INTRODUCTION

Numerous studies have attempted to understand the factors that make some individuals more or less risk averse in building their investment portfolios, whether in "normal" times or in times of economic downturns. However, none of those studies have utilized the recently published Zimbardo Time Perspective Index (ZTPI) created by Philip Zimbardo of Stanford and John Boyd

of Google (Zimbardo and Boyd, 2008), which is very persuasive in postulating that one's perspectives concerning past, present, and future have a major impact on financial decisions. The primary purpose of this study is to examine the effects of how an individual's perception of time determines that individual's attitudes and intended actions concerning job changes in times of national/global financial crises. The secondary purpose of this study is to test whether the standard scale items are more applicable than tailored scale items to this type of problem.

Specifically, the study primarily seeks to learn whether any particular *time perspectives* have strong effects on an individual's intention to change jobs during this financial crisis. The timing was initially fortuitous for researchers, providing a large sample of respondents who were coping with major financial concerns. A secondary objective of this study is to determine whether a second set of scale items that includes questions that are more focused on attitudes toward finances may be a useful refinement of the ZTPI for financial decisions.

This study first analyzed the responses to questionnaires distributed in the fall of 2008 through the spring of 2009, by which time most households, or their friends and extended family members, were feeling some effects of the worst worldwide recession since the Great Depression. Median household net worth fell nearly 40% between 2007 and 2010 (Riley, 2012). Corporations reacted too: of cash grew to nearly two trillion dollars of cash, dubbed "scared money" by 2011 (Whitehouse, 2011). This initial study used the standard ZTPI and other questions designed to capture how economic changes affect one's tendency to seek new employment and his/her intention to move assets to safer forms of investment.

Then, in 2010, when the economy started showing strong signs of recovery – *except for growing unemployment* – a second set of scale items was tested on different respondents, using the same questions from the first model but also including 21 additional questions. These additional questions are based on the ZTPI questions, but their re-wording was designed to focus on time orientation relative to finances. It was expected that these additional questions would provide a refinement for the purpose of this study, and thus would have more explanatory power regarding financial decisions.

One more time, in 2012, when unemployment had leveled off and begun to decline, and as consumer confidence was on the rise, more data were gathered for a further revised model, using 60 questions tailored towards individuals' TPIs applied toward financial issues, including some version of the revised 21 questions distributed in 2010.

LITERATURE REVIEW

Previous studies have advanced our understanding of why some people change jobs and/or careers. Murtagh, Lopes and Lyons' (2012) study supports other-than-rational perspectives of career decision making. Chambers, Benibo and Spencer (2011) examined the usefulness of the theory of planned behavior for explaining the actions individuals intend to take concerning (1) moving funds and (2) changing jobs during a financial crisis. That study's results indicate that the theory of planned behavior substantially explains the intent to move funds but is only moderately useful in predicting job changes.

Time perspective as an explanation for behavior has a long history in the psychology literature. Raju (1980) posited that past-oriented shoppers tend to be rigid and more risk averse, and this is why they tend to not buy on impulse. Gonzalez and Zimbardo (1985) indicated that

future-oriented individuals characteristically delay gratification. The future-oriented individual is also less likely to take risks (Lennings and Burns, 1998; Zimbardo and Boyd, 1999). According to the findings of Karande and Merchant (2012), past orientation positively affects prudence but has no significant effect on impulsiveness; Present orientation has a positive effect on impulsiveness and a negative effect on prudence – both as expected; and future orientation positively affects prudence but has no significant effect on impulsiveness.

Time perspective "is often the non-conscious personal attitudes that each of us holds toward time and the process whereby the continual flow of existence is bundled into time categories that help to give order, coherence and meaning to our lives" (Zimbardo and Boyd, 2008:51). They contend that individuals vary significantly in terms of time perspective, and that these differences strongly and predictably influence individuals' responses to particular situations. They specifically identified six orthogonal perspectives, namely: past-positive, past-negative, present-fatalistic, present-hedonistic, future, and transcendental-future. Some individuals are strongly oriented to one or more of these perspectives while moderately or weakly oriented to the others. Below is a brief explanation of the expected beliefs and actions for a strong orientation toward each of these perspectives:

Past-positive: The past-positive time perspective yields a favorable interpretation of and attitude toward past events, even when those events may objectively be negative. To the extent that "what people *believe* happened in the past influences their present thoughts, feelings and behavior more than what really happened" (Zimbardo, 1980:61), people with a strong past-positive time perspective are not likely to be discouraged by past financial loss. Indeed, they might, like an athlete, see it as the necessary "pain" before a "gain."

Past-negative: On the contrary, an individual with a past-negative time perspective will reflect on the past with thoughts of "I could have done better," being predisposed to wondering if it was necessary to go through "pain" in order to have a "gain." A person with this time perspective is likely to take the least risky financial actions, especially if having already experienced less-than-expected returns or a loss on personal investments.

Present-hedonistic: The individual with a strong present-hedonistic perspective is one who is driven by instant gratification: If it feels good, do it is her/his philosophy (Zimbardo and Boyd, 2008). This means that future consequences (positive or negative) do not direct present behaviors. The strong present hedonist is not only unlikely to have any investment plan but would also not reduce spending on non-essentials even in a declining economy unless directly affected, as with personal job loss. The past, bad or good, is considered as gone forever and therefore undeserving of regrets. Having saved money at all may depend on a low present-hedonistic outlook. Lusardi (1999) finds that households consume today at the expense of tomorrow when they lack self-control (high present-hedonistic) and have not planned for retirement (low future orientation).

Present-fatalistic: The behavioral responses to a declining economy of a person with a strong present-negative time perspective would be similar to those of the preceding type, but the rationale is different. The present fatalist lives for the moment because s/he perceives forces beyond her/his control determining the outcome. Consequently, the strongly present-fatalistic person not only refuses to make plans for the future, but is willing to live with whatever happens.

Future: In contrast to both the strong present hedonist and strong present fatalist, the person with a strong future time perspective is "conscientious, consistent, and concerned with future

consequences" (Zimbardo and Boyd, 2008:64). The future-oriented individual makes very deliberate plans toward desired goals and always conscious of how s/he spends time and is likely to decrease spending in a declining economy as well as take rational steps to protect investments from future losses.

Transcendental-future: This is the most recently identified of these six perspectives: this individual does not view death as the end of life. Material aspects of life, including personal finances and investments, are considered to be transient concerns, not affecting one's ultimate, eternal goals. The individual with a strong transcendental-future orientation continues to feel hope, even under the most dire circumstances. We do not test for this perspective in our models.

Although Zimbardo and Boyd (2008) contend that time orientation is acquired via socialization both at cultural and sub-cultural levels, its effects on subsequent attitudes and behaviors are "non-conscious." One's time orientation has ontological as well as epistemological consequences beyond one's deliberate control. Hence, an individual who, for example, has as a cyclical view of time, as opposed to a linear view, habitually arrives late for appointments and procrastinates in spite of a "conscious" dislike of these behaviors.

Of course, an individual can experience any of these six perspectives at various times. The ZTPI provides a scale that indicates which of these six are strong (with relatively high Index numbers) and are therefore likely to be the individual's major ways of perceiving, vs. those which are weak and have relatively little influence over thoughts and actions.

HYPOTHESES

The ZTPI is the foundation for five hypotheses concerning intent to change jobs. This study departs from the ZTPI in two major ways: (1) After the original ZTPI scale was used in 2008 and 2009, producing marginally significant but not robust results against both intent to change jobs and intent to move assets out of stocks and bonds and into cash, 21 new, financially-tailored ZTPI-type questions were added to the instrument, mainly for past-positive, present-hedonistic and future perspectives, to see if scales that are more financially focused would more strongly capture attitudes toward finances. Because the stock market had begun to stabilize by 2010, intent to move investments into cash is no longer expected to be significant. In 2012, the 21 items were refined further and these constructs were developed and tested, for a total of 60 items. It was expected that individuals would still be *feeling* the effects of the economic crisis, although for most families, the crisis was in the past.

The ZTPI, which is a general scale on one's outlook on life, may produce unexpected results when applied to this or any other specific situation. For that reason, the second and third survey instruments include both the original scale and the financially-tailored questions, in order to collect information beyond that pertaining to the general ZTPI. The first survey instrument is shown in Appendix A. The results of this survey are modeled in Figure B1, shown in Appendix B. These results show that tailoring the scales used to measure the ZTPI constructs may be useful. Only the 2010 survey results, which test whether tailoring is useful against the original ZTPI scale items, and the 2012 survey results with further refined financial scales, are discussed here. The 21 new time perspective financial constructs utilized in 2010 (questions 78 through 101 and later referred to as TPIF items) are shown in Appendix C, and the 60 TPIF2 items tested in 2012 are shown in Appendix D. Some of the survey items are reverse-scaled to protect against positive

response bias. Written instructions were included with the instrument to the participants, as was an assurance of the confidentiality of participants and an explanation of the voluntary nature of participation.

Dependent Variable: Intent to Change Jobs

Intent is the extent to which a person expresses a willingness to exert effort in order to perform the specified behavior, e.g., changing jobs. In this paper, intent to react to the effect of the national/global financial crisis on one's own financial security is measured by changing income streams, with a voluntary employment change. For example, respondents were asked on a 5-point scale how true - 1 = very untrue to 5 = very true - was the following statement: "As a result of how I feel now, I intend to look for a new job."

How a Strong Time Perspective Orientation Is Expected to Influence Intent

A strong past-positive orientation means that a person has a favorable interpretation of and attitude toward past events, even if those events may objectively be negative. The more positive one's perception of the past, the more faith s/he may have that things will all work out well and the less likely s/he is to intend to reallocate assets or change jobs:

H_1: A strong past-positive orientation will have a *negative* effect on the intent to change jobs.

This and the subsequent relationships are shown pictorially as arrows from the dependent variable to the independent variable, for all hypotheses, in Figure 1.

A strong past-negative orientation means that a person has an unfavorable interpretation of and attitude toward past events, even if those events may objectively be positive. Therefore, never satisfied, we expect those with a strong past-negative orientation will more readily consider changing jobs:

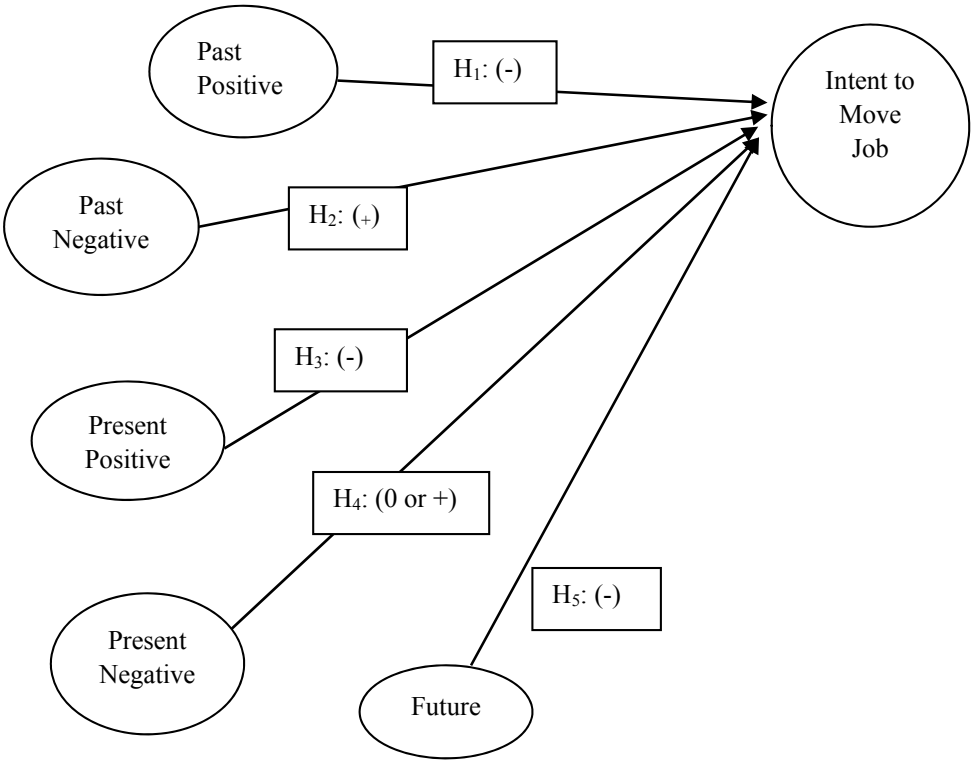
H_2: A strong past-negative orientation will have a *positiv*e effect on intent to change jobs.

An individual with a strong present-positive, hedonistic orientation lives for pleasure in the moment. The present hedonist is not only unlikely to have an investment plan but would also not reduce spending on non-essentials even in a declining economy. Therefore, people who are living for today may intend but would not take the trouble to voluntarily change jobs, which is a tedious and stressful process:

H_3: A strong present-hedonistic orientation will have a *negative* effect on the intent to change jobs.

The person with a strong present-negative, fatalistic orientation actually perceives her/himself to be suffering from this economy. The hypothesis here is that pain avoidance will dominate the fatalism. Therefore, if a person perceives sufficient control to change jobs, then the relationship between present-fatalism and intent to change jobs will be positive. However, that relationship is expected to be weak.

Figure 1 – 2010 and 2012 Research Models



H4: A strong present-fatalistic orientation will have either *no* effect or a weakly *positive* effect on the intent to change jobs.

The individual with a strong future orientation is concerned with future consequences of present actions. That is, the future oriented individual feels responsible for taking actions that will lead toward goal attainment. But, taking a long-term view of situations, future-oriented people are expected to be more likely to view the economic turbulence as temporary and therefore are more likely to stay the course and are less likely to intend to change jobs:

H5: A strong future orientation is expected to have a *negative* effect on intent to intent to change jobs.

Several demographic control variables - including age, gender, household income, racial identity, religiosity and experience - were also tested, with no significant results expected. Taken as a whole, the model can be expressed pictorially, as shown in Figure 1.

METHODOLOGY

The 2008/2009 Model

Initially, data were gathered in late 2008 through early 2009, during the worst of the financial crisis, measuring the generic ZTPI items, and those constructs' effect on two dependent variables: The first dependent variable, "Intent to Move Jobs," is operationalized as a composite of participants' responses to survey questions. Respondents were asked to indicate, for example, on a 5-point scale how true (1 = very untrue to 5 = very true) was the following statement: "As a result of how I feel now, I intend to look for a new job." The second dependent variable, "Intent to Reallocate Assets," is also operationalized as a composite of participants' plans, in this case, concerning two aspects of their intent to move their assets from financial markets: (1) to banks, and (2) to cash. Again, respondents were asked to indicate, for example, on a 5-point scale how true (1 = very untrue and 5 = very true) were the following statements: "…I intend to move my financial assets from financial markets to cash" or "… I intend to move my financial assets from financial markets into banks." The results of this integrated model were promising, but weak, resulting in two subsequent revisions of scale items. Intent to Reallocate Assets was also dropped from subsequent models, due to the subsequent stabilization of financial markets. The 2008/2009 results are presented in Appendix B and not discussed further here.

2010 Model: Adapting the 2008/2009 Independent Variable Scale Items

The primary independent variables for the 2008/2009 model are the ZTPI constructs from Zimbardo and Boyd's (2008) scale, excluding those for transcendent future orientation. This model produced weak but promising results. This led the authors to question, if items were constructed to be more specific to people's time perspectives on finances, would survey results be more robust? That is, can adapted TPI scales create a more predictive model of people's reactions to a national economic crisis? To test for this research question, scale items that are more focused on financial issues were created to measure the five original ZTPI constructs. With the inclusion of these more focused scale items, it is expected that the revised model will explain more of the variance in intent to move jobs. However, since the data for the 2010 Model were gathered in late 2010, those who intended to move investments into more stable vehicles probably had made that move by then; and with the financial markets stabilizing, intent to move money was no longer expected to be significant. This second, 2010 instrument includes the original ZTPI variables along with 21tailored scale items based on the original ZTPI scale items but written specifically for measuring attitudes toward finances (referred to as TPIF). The tailored scale items are shown in Appendix C.

The questions from the original ZTPI scale are numbered in Appendix A, and the TPIF scale in Appendix C, as:

- Past-negative question numbers: 4, 5, 16, 22, 27, 33, 34, 36, 50, 54in the ZTPI.

- Past-positive: 2, 7, 11, 15, 20, 25, 29, 41, 49 in the ZTPI and #84 only constructed for the 2010 TPIF,

- Present-fatalistic: 3, 14, 35, 37, 38, 39, 47, 52, 53 in the ZTPI, and 89, 94 and 95 constructed for the 2010 TPIF.

- Present-hedonistic: 1, 8, 12, 17, 19, 23, 26, 28, 31, 32, 42, 44, 46, 48, 55in the ZTPI and 81,83, 88, 90, 92, 93, 96 and 98 constructed for the 2010 TPIF.

- Future: 6, 9, 10, 13, 18, 21, 30, 40, 43, 45, 51, 56in the ZTPI and 82, 85, 86, 87, 91, 97, 99, 100 and 101 constructed for the 2010 TPIF.

Demographic control variables examined are gender, age, race, ethnicity, religiosity, assets and monthly household income. Age is measured in terms of how old the respondent was at her/his last birthday. Monthly household income is measured in terms of the respondent's recollection of the approximate total of all household income earned. Gender is dummy-coded 0 = male, and 1 = female. Race and ethnicity are divided into the two ethnic groups large enough to analyze in this group of respondents, and "other." For the "white" construct, participants are coded 1 = white and 0 for other; for the Hispanic construct, participants are coded 1 = Hispanic and 0 for other. (No other non-white, non-Hispanic group was large enough in this sample to be analyzable.) Similarly, given the small amount of variation present in this sample, religious affiliation is dummy coded 0 = Catholics, 1 = Non Catholics.

2012 Scales: Refining the 2010 Independent Variable Scale Items

While the 2010 model produced materially better results than the 2008/2009 model, there was room for improvement in the convergent validity of the scales. Thus, several new specific items were tested using only the revised tailored scale (TPIF2) in late 2012. The model was also analyzed to evaluate the persistence of the 2010 findings.

The questions from the TPIF2 scale utilized in 2012 are numbered in Appendix D as:

- Past-negative question numbers: 4, 16, 26, 33, 49 and 53.

- Past-positive: 11,19,24 and 40,

- Present-fatalistic: 36, 37, 38 and 46.

- Present-hedonistic: 8, 12, 22, 41, 43, 45 and 54.

- Future: 10, 39, 42, 44 and 50.

Sample and Data Collection

2010 Model. Approximately 221 members of a South Texas university's students participated in the 2010 version of this survey, producing 220 usable responses. Investigators asked colleagues for permission to survey their students, who ranged from freshmen through graduates. The median age of the participants was 23 years, with a range from 17 to 53. Fifty-three percent of the participants are female. Respondents' experience varied from very little perceived business

experience to extensive business experience, with the average participant's self-rating as an experience level of 2.8 on a 5-point scale.

2012 Model. An additional 495 students were surveyed for the 2012 version of this study. The median age of the participants was again 23 years, with a range from 16 to 68. Fifty-five percent of the participants are female. Respondents' business experience also showed great variation, with the average participant's self-rating as an experience level of 2.6 on a 5-point scale. Students were selected on the basis of convenience. Care was taken, nevertheless, to ensure that participating students were distributed from freshmen through graduates.

ANALYSIS

To explore any possible bias resulting from the use of students, bivariate correlations were examined between demographic variables and dependent variables of interest. No significant correlations were found, except as noted in the results section. Very few of the respondents were 18 or under, and removing these responses from our data set did not significantly affect our results. Based on these results, it appears that demographic factors are generally not significant in explaining intent; therefore, the use of student subjects, whose demographic data may not be reflective of the general population, can provide useful information.

The overall model and the individual scales used to measure constructs and their underlying latent constructs, shown in Figure 1, are assessed using partial least squares (PLS) analysis. PLS is the most logical method to analyze the theoretical model because it simultaneously addresses both the effectiveness of the model and the reliability of the general underlying measures as applied to this specific economic study. Other advantages of PLS are relaxed error and distribution assumptions (Wold, 1982).

PLS factor loadings are calculated to assess the construct validity of each of the measurement items. A factor loading of 0.707 or greater is considered to be a substantial correlation between the indicator and the latent variable (Barclay et al., 1995; Chin, 1998); however Barclay et al. (1995) note that it is not uncommon for items in newly developed scales to fail to meet the 0.707 level of reliability. Where factor loadings are smaller, these items will generally be weighted less, because PLS minimizes the error variance for the whole model.

RESULTS

Both the 2008/2009 Model (shown in Appendix B) and the 2010 Model were analyzed similarly. The analysis techniques for the 2008/2009 model and the results are relegated to Appendix B for parsimony. For 2010, the results are discussed in detail. The results from the 2012 revised scales Model are contrasted to the 2010 Model here.

Comparing ZTPI Scale Items to TPIF Items in the 2010 Model

Because the underlying scales, developed for day-to-day decisions and actions, do not capture the essence of such financial attitudes and choices under financial duress, the 2010 Model

was run with the inclusion of 21 modified present-hedonistic and future items to see if they would be better for capturing the impact of more specific attitudes on intent regarding finances.

The results of the confirmatory factor analysis suggest that the measurement items within each original scale are only moderately correlated with the underlying latent variable. The factor loadings and weights for data collected in 2010 are for the ZTPI items is shown in Table 1, and Table 2 shows results for the TPIF 2010 items.

Table 1. 2010 Model, ZTPI Measurement Variables, Under Partial Least Squares

Factor	Item #	Factor Loading	Weight	Factor	Item #	Factor Loading	Weight
Intent – Move Job	74	0.8985	0.4473				
Intent – Move Job	75	0.7982	0.3219				
Intent – Move Job	76	0.8301	0.4110				
Present Hedonistic	1	0.0250	0.2019	Past Positive	2	0.3404	0.0921
Present Hedonistic	8	-0.0463	-0.0452	Past Positive	7	0.2439	0.2490
Present Hedonistic	9	-0.0460	-0.0313	Past Positive	11	0.4621	0.4438
Present Hedonistic	12	-0.0722	-0.0391	Past Positive	15	-0.1121	-0.0918
Present Hedonistic	17	0.0798	0.0548	Past Positive	20	-0.0672	-0.0506
Present Hedonistic	19	-0.0501	-0.0351	Past Positive	29	0.2342	0.0994
Present Hedonistic	23	-0.0420	-0.0293	Past Positive	49	0.8222	0.8099
Present Hedonistic	24	-0.0308	-0.0392	Past Negative	4	0.1509	0.0705
Present Hedonistic	26	0.7849	0.4701	Past Negative	5	-0.1140	-0.0618
Present Hedonistic	28	0.7713	0.3947	Past Negative	16	-0.0763	-0.0671
Present Hedonistic	31	0.3918	0.2529	Past Negative	22	-0.1160	-0.1035
Present Hedonistic	32	-0.0020	-0.0179	Past Negative	25	-0.0534	-0.0498
Present Hedonistic	42	-0.0075	-0.0186	Past Negative	27	0.6801	0.6492
Present Hedonistic	44	0.3990	0.2834	Past Negative	33	0.6793	0.6625
Present Hedonistic	46	0.2913	0.2866	Past Negative	34	-0.0292	-0.0294
Present Hedonistic	48	0.1643	0.0847	Past Negative	36	-0.1213	-0.0685
Present Hedonistic	55	0.0427	-0.0119	Past Negative	50	0.2956	0.1862
Present Hedonistic	56	0.0520	0.0358	Past Negative	54	-0.0608	-0.1115
Present Fatalistic	3	0.5373	0.2567	Future	6	0.2012	0.1940
Present Fatalistic	14	0.4925	0.1112	Future	10	0.1808	-0.1495
Present Fatalistic	35	0.5422	0.5374	Future	13	-0.3396	-0.4323
Present Fatalistic	37	0.4277	-0.0187	Future	18	0.5468	0.5821
Present Fatalistic	38	0.5589	0.2655	Future	21	0.1633	0.1768
Present Fatalistic	39	0.5568	0.3662	Future	30	0.4732	0.5767
Present Fatalistic	41	0.0163	-0.0397	Future	40	0.1418	0.1260
Present Fatalistic	47	-0.1394	-0.1067	Future	43	0.5304	0.4018
Present Fatalistic	52	0.4795	0.2919	Future	45	-0.0177	-0.1286
Present Fatalistic	53	-0.1412	-0.1232	Future	51	0.1550	-0.0788

Less than half of the average variance for each factor is explained, with the exception of the intent items. This indicates that the measurement items in these scales exhibits only moderate convergent validity and are highly correlated to each other due to a single underlying construct. Chin (1998) and Höck and Ringle (2006) considered a composite reliability of 0.6 or greater to be adequate for an exploratory model. The average variance explained substantially increased with the TPIF scales, and composite reliability also materially increased, indicating that some tailoring

of scale items is an improvement over the general scale items in this case. Except for the past negative scale (which was *not* tailored), each of the TPIF reliability statistics exceeded the recommended value of 0.60.

Table 2. 2010 Model, TPIF Measurement Variables Using Partial Least Squares

Factor	Item #	Factor Loading	Weight	Factor	Item #	Factor Loading	Weight
Intent – Move Job	74	0.8982	0.4434	Past Negative	4	0.1464	0.0683
Intent – Move Job	75	0.8088	0.3422	Past Negative	5	-0.1155	-0.0623
Intent – Move Job	76	0.8217	0.3955	Past Negative	16	-0.0775	-0.0677
Present Hedonist New	81	0.9710	0.1626	Past Negative	22	-0.1160	-0.1041
Present Hedonist New	83	0.9744	0.1653	Past Negative	25	-0.0551	-0.0509
Present Hedonist New	88	0.8397	0.1497	Past Negative	27	0.6802	0.6500
Present Hedonist New	90	0.9704	0.1590	Past Negative	33	0.6796	0.6630
Present Hedonist New	92	0.7376	0.1146	Past Negative	34	-0.0311	-0.0298
Present Hedonist New	93	0.8366	0.1255	Past Negative	36	-0.1232	-0.0690
Present Hedonist New	96	0.7036	0.1120	Past Negative	50	0.2902	0.1824
Present Hedonist New	98	0.7921	0.1675	Past Negative	54	-0.0660	-0.1141
Present Fatalist New	89	0.8033	0.3639	*Future New*	82	0.9544	0.1600
Present Fatalist New	94	0.9088	0.3021	*Future New*	85	0.5989	0.0728
Present Fatalist New	95	0.8975	0.4825	*Future New*	86	0.9605	0.1542
Past Positive New	84	1.0000	1.0000	*Future New*	87	0.8211	0.1195
				Future New	91	0.9589	0.1525
				Future New	97	0.9302	0.1156
				Future New	99	0.9283	0.1189
				Future New	100	0.9296	0.1207
				Future New	101	0.8268	0.0979

Table 3 compares the average variance explained and the composite reliability for both the 2010 ZTPI items and the 2010 TPIF items. The TPIF, financially-tailored scales explained more variance and were more reliable than the ZTPI measures.

Table 3. Two-Model Comparison of Common Variance Explained and Composite Reliability Measures, 2010 Data

Construct	Average Variance Explained, General ZTPI Items	Average Variance Explained, TPIF Items	Composite Reliability, ZTPI Items	Composite Reliability, TPIF Items
Intent – Change Jobs	0.711	0.712	0.881	0.881
Past Positive*	0.162	1.000	0.470	1.000
Past Negative**	0.099	0.099	0.363	0.364
Present Hedonistic	0.092	0.738	0.400	0.957
Present Fatalistic	0.190	0.759	0.652	0.904
Future	0.106	0.785	0.458	0.970

The correlations among these tailored latent variables are presented in Table 4; the numbers presented in the diagonal depict the square root of the average common variance extracted by the

measurement items within the scale (the average inter-item correlation). The correlations among the latent variables are smaller than the square root of the common variance extracted within each scale - with the exceptions of the future TPIF, which correlates more highly with both the new constructs for present-hedonistic and present-fatalistic – demonstrating an overall divergent validity where the items within a scale are more significantly related to one another than to items in other scales. Present-hedonistic and present-fatalistic are also highly correlated with one another. Based on the preceding results, the measurements exhibit moderate validity and reliability, but could be improved.

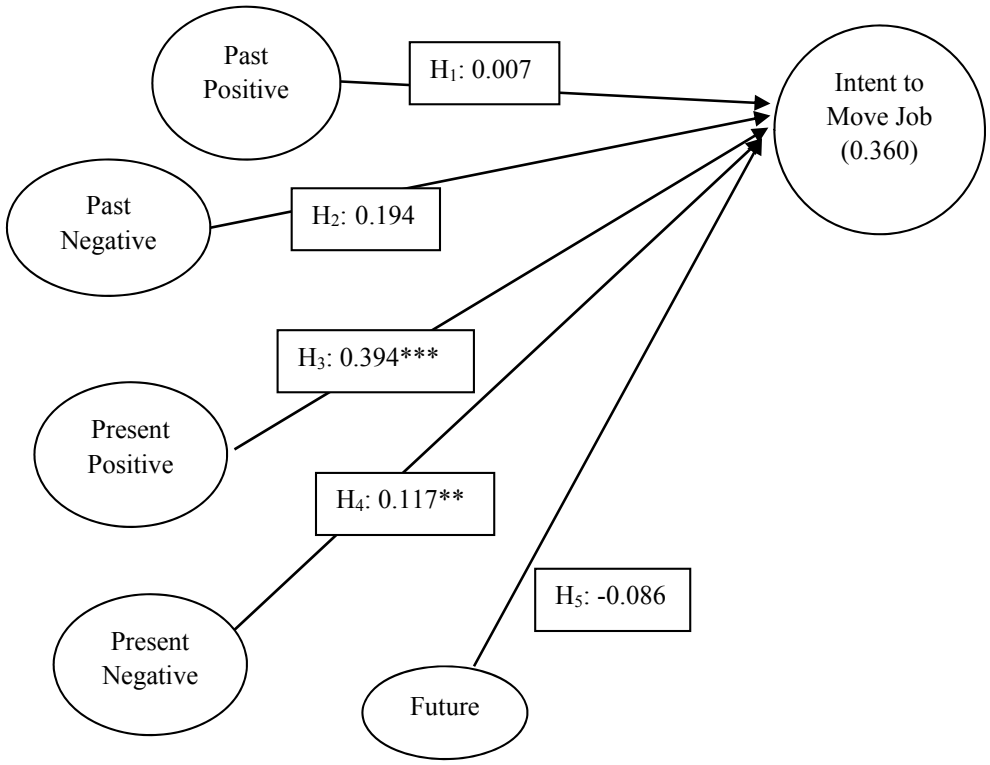
Table 4. 2010 TPIF Item Model Correlations among Latent Variables

Construct	Intent to Change Job	*New* Past Positive	Past Negative	*New* Present Hedonistic	*New* Present Fatalistic	*New* Future
Intent to Change Job	**0.844**					
New Past Positive	0.217	**1.000**				
Past Negative	0.500	-0.031	**0.315**			
New Present Hedonistic	0.374	0.686	-0.021	**0.859**		
New Present Fatalistic	0.409	0.586	0.179	0.889	**0.871**	
New Future	0.320	0.673	-0.037	0.968	0.904	**0.886**

The path coefficients to the dependent variables from the latent variables are presented in Figures 2 and 3. In the General Model, two path coefficients are significant at $p \leq 0.05$, but only one of those paths is above 0.20: Recalling the hypotheses:

H_3: A strong present-hedonistic orientation will have a *negative* effect on the intent to change jobs.

Figure 2 – 2010 ZTPI Model Using Partial Least Squares



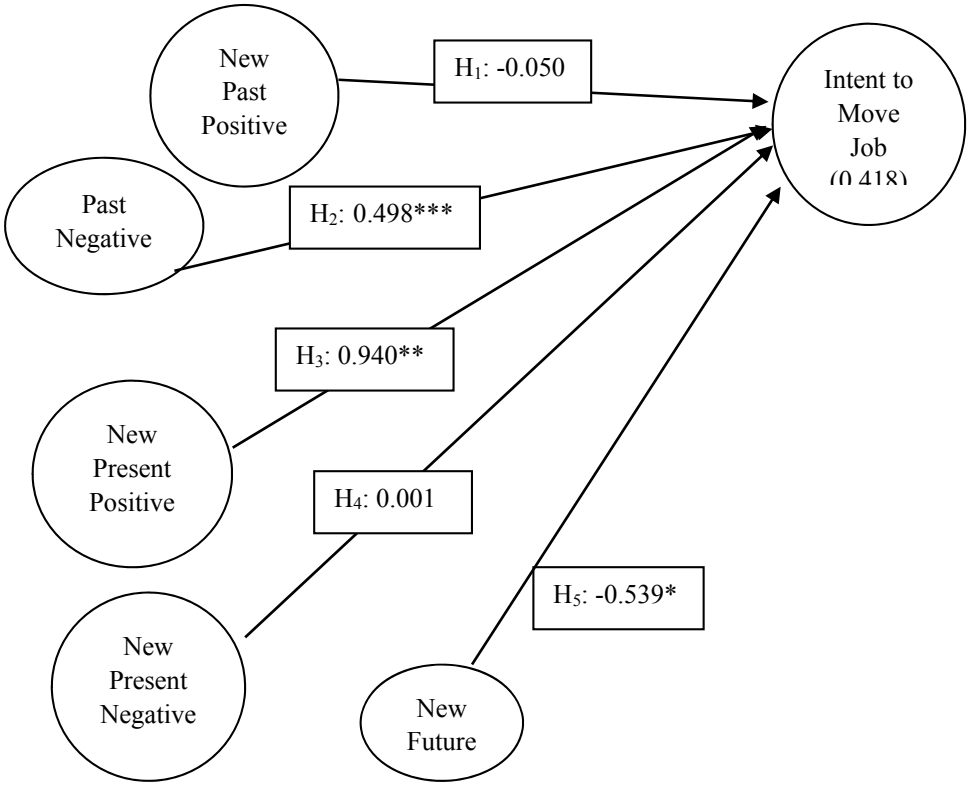
In the TPIF Model, two path coefficients are significant at $p \leq 0.05$, and a third is marginally significant at $p \leq 0.10$ and all of those paths are above 0.20:

H_2: A strong past-negative orientation will have a *positiv*e effect on intent to change jobs.

H_3: A strong present-hedonistic orientation will have a *negative* effect on the intent to change jobs.

H_5: A strong future orientation is expected to have a *negative* effect on intent to intent to change jobs.

Figure 3 – 2010 Model, TPIF Structural Model Using Partial Least Squares



Refining the Scales and using 2012 Data

While the TPIF scales worked better than the ZTPI scales in 2010, the researchers postulated that by improving the TPIF scales, more path coefficients would be significant. Consequently, several new items were developed and tested, as shown in Appendix D. The best scale items from 2012 overall are conceptually aligned with the ZTPI scale, shown in Table 5.

Comparing the two scales shows mixed results. (See Table 6, where the preferred scale results are highlighted.) Both Past-Positive and Past-Negative scales are improved, but the original Present Hedonistic, Present Fatalistic and Future scales were not. While the composite reliability is adequate under both scales, the average variance explained in those scales dropped below 0.50, indicating a marked worsening in those scales. Further, while the average variance explained (AVE) was improved for both Past Positive and Past Negative scales, the refinement process should continue until those scale AVEs at least equal 0.50. Note that the Intent scale did not change across years, and differences in average variance explained and composite reliability in scale results between the years are negligible.

Table 5. 2012 TPIF2 Measurement Variables Using Partial Least Squares

Factor	Item #	Factor Loading	Weight	Factor	Item #	Factor Loading	Weight
Intent – Move Job	58	0.8026	0.4071	*Past Negative*	4	0.0357	0.0308
Intent – Move Job	59	0.7803	0.3990	*Past Negative*	16	0.2561	0.1650
Intent – Move Job	60	0.8293	0.4364	*Past Negative*	26	0.2417	0.0668
Present Hedonist	8	-0.0653	-0.0382	*Past Negative*	33	0.7404	0.4282
Present Hedonist	12	0.1727	0.0848	*Past Negative*	49	0.7095	0.4447
Present Hedonist	22	0.3683	0.1584	*Past Negative*	53	0.7183	0.4286
Present Hedonist	41	0.6810	0.2982	*Future*	10	0.2808	0.1133
Present Hedonist	43	0.6751	0.3229	*Future*	39	0.5384	0.2233
Present Hedonist	45	0.7096	0.3865	*Future*	42	0.6787	0.3430
Present Hedonist	54	0.6803	0.3368	*Future*	44	0.7943	0.3990
Present Fatalist	36	0.6259	0.3372	*Future*	50	0.7820	0.3813
Present Fatalist	37	0.6648	0.3447	*Past Positive*	11	0.2430	0.0342
Present Fatalist	38	0.7429	0.3901	*Past Positive*	19	0.5792	0.3552
Present Fatalist	46	0.6652	0.4059	*Past Positive*	24	0.2730	0.0874
				Past Positive	40	0.9191	0.8292

Table 6. Two-Scale Comparison of Common Variance Explained and Composite Reliability Measures, TPIF Items, 2010 and TPIF2 Items, 2012

Construct	Average Variance Explained, 2010 Items	Average Variance Explained, 2012 Items	Composite Reliability, 2010 Items	Composite Reliability, 2012 Items
Intent – Change Jobs	0.712	0.647	0.881	0.846
Past Positive	1.000	0.382	1.000	0.602
Past Negative	0.099	0.282	0.364	0.629
Present Hedonistic	0.738	0.294	0.957	0.694
Present Fatalistic	0.759	0.457	0.904	0.770
Future	0.785	0.414	0.970	0.763

Evaluating the correlations among latent variables in Table 7, the 2010 correlations indicated problems among Present Hedonistic, Present Fatalistic and Future constructs, yet the 2012 scales only moderately corrected some of those problems, and new but mild problems with the Past Negative construct arose. Troublesome correlations are highlighted in Table 7.

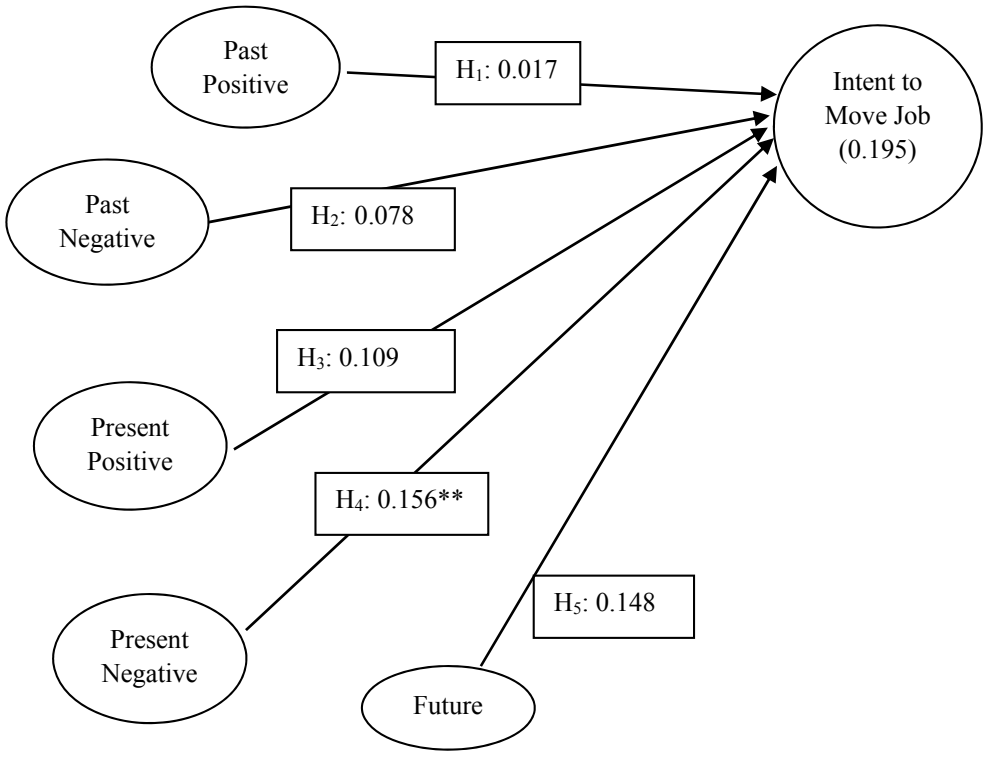
Table 7. 2012 TPIF2 Item Model Correlations among Latent Variables

Construct	Intent to Change Job	Past Positive	Past Negative	Present Hedonistic	Present Fatalistic	Future
Intent to Change Job	**0.804**					
Past Positive	0.210	**0.572**				
Past Negative	0.318	0.489	**0.531**			
Present Hedonistic	0.354	0.402	0.555	**0.542**		
Present Fatalistic	0.360	0.356	0.508	0.605	**0.676**	
Future	0.371	0.341	0.565	0.670	0.341	**0.643**

* The numbers presented in the diagonal depicting the square root of the average common variance extracted by the measurement items within the scale.

With the 2012 data, one path coefficient is significant at $p \leq 0.05$, but that path is not above 0.20. Further, the overall R-squared for the model dropped to 0.195. See Figure 4.

Figure 4 – 2012 TPIF2 Results Using Partial Least Squares



DISCUSSION

The ZTPI and the two models with scales that focused on finances have results that suggested time perspective does indeed capture some predictable differences among individuals. A comparison of the models is found in Table 8.

Table 8 -- Comparison of ZTPI, TPIF and TPIF2 Models

		ZTPI Model		2010 TPIF Model		2012 TPIF2Model	
Hypothesis	Expected Sign	Actual Path	Significant?	Actual Path	Significant?	Actual Path	Significant?
1	(-)	0.007	N	-0.050	N	0.017	N
2	(+)	0.194	N	0.498	***	0.078	N
3	(-)	0.394	***	0.940	**	0.109	N
4	0	0.117	**	0.001	N	0.156	**
5	(-)	-0.086	N	-0.539	*	0.148	N

*** is significant at p<.01; ** is significant at p<.05; * is marginally significant at p<.10.

The 2010 ZTPI Model had an R-squared of 0.360 and only one significant path size larger than 0.200, and that path had an unexpected sign. By comparison, the 2010 TPIF Model had an R-squared of 0.418 and three path sizes larger than 0.200. Two of those path sizes were significant and one was marginally significant, although one of the two significant paths again had an unexpected sign. The 2010 TPIF Model appears to be an improvement over the 2010 ZTPI Model, although both are reasonably robust predictors of intent to move jobs in a crisis. The 2012 TPIF2 results were not an especially good predictor; by then, unemployment had not materially improved but had stabilized and was decreasing. Understandably, the economic crisis appears to have prompted an emotional, "scaredy cat" response, with people relying more on their own intuition as colored by their individual time perspectives, rather than relying on external, more rational, "cool cat" factors that one might use in more "normal" times.

In the 2010 TPIF Model, H_2 (Past Negative to Intent) is significant with a sufficient path size. Similarly, H_3, (Present Hedonistic to Intent) is significant with a robust path size under the TPIF Model; but in both models, the sign is positive instead of negative. Although this was not expected, it would make sense if cautious people saw the tenuous economic situation as a bad time to take a risk on a new job, but a carefree person did not see any danger in forging ahead. H_4, Present Fatalistic, is very close to zero, as predicted. H_5, (Future to Intent) is negative, as predicted, and marginally significant in the TPIF Model but insignificant in the ZTPI Model.

Still, further examination is in order. The scales are improving but need more improvement and further testing in times of crisis. The reasons why and at what point a person begins to rely heavily on time perspective in making a job or career decision, and whether the shift to greater reliance on time perspectives happens suddenly or very gradually, require further study. Although one does

not wish for an economic downturn or other crisis to study the shift to "scaredy-cat" decision-making, that appears to be the best opportunity for continuing this line of research.

LIMITATIONS AND FUTURE DIRECTIONS

This study was limited to examining individual responses to a general financial crisis, with three samples, over a four year period. The first iteration focused on only two issues related to positioning oneself for a better financial outcome in response to a major economic downturn – jobs and investing funds in safer vehicles, at a time when many of the financial experts did not have much advice to give. The second and third iterations focused only on changing jobs. The marked improvement in the economy over this four-year time period was an impediment to measuring the improvement in the survey questions, as individual decision-making appears to have moved from "scaredy cat" mode to a more normal "cool cat" mode. In order to not wait for another recession to test these measures, it might be useful to survey individuals' response to a more localized financial downturn, or a localized natural event, that affects employment or the value of assets.

CONCLUSION

The use of the ZTPI for examining a myriad of financial decisions – whether personal or acting within a firm – may hold great promise for better understanding and greater predictability of those decisions. However, this study shows that adapting the scales to a greater level of specificity to the dependent variable of interest adds predictive value, at least in this case.

Further, peoples' reliance on internal time perspectives appears to be greater in times of crisis than in more "normal" times. During crises, the stronger one's Past Negative and/or Present Hedonistic perspective, the more likely one is to change jobs. Conversely, Futurists wait out crises and are less likely to change jobs. Additional studies under differing circumstances can shed more light on the general usefulness of time perspective in this realm.

REFERENCES

Barber, B. M. and T. Odean. (2001). Boys will be boys: gender, overconfidence and common stock investment. *Quarterly Journal of Economics*, *116*(1), 261-292.

Barclay, D., R. Thompson, and C. Higgins. (1995). The partial least squares (pls) approach to causal modeling: personal computer adoption and use as an illustration. *Technology Studies*, 285-323.

Chambers, V., B. Benibo, and M. K. Spencer. (2011), Reactions to the 2008 economic crisis and the theory of planned behavior. *Academy of Accounting and Financial Studies Journal.*

Chin, W.W. (1998). The partial least squares approach for structural equation modeling. (1998). *Modern methods for business research*. Mahwah, NJ: Lawrence Erlbaum Associates.

Chin, W.W., and P. R. Newsted. (1999). Structural equation modeling analysis with small samples using partial least squares, In R. H. Hoyle, ed. *Statistical Strategies for Small Sample Research*, Thousand Oaks, CA: Sage Publications.

Gonzalez, A., and P. G. Zimbardo (1985). Time in perspective: the time sense we learn affects how we do our jobs and enjoy our pleasures. *Psychology Today,* 19(3), 21–16.

Höck, M., and C. M. Ringle. (2006). Strategic networks in the software industry: an empirical analysis of the value continuum. IFSAM VIIIth World Congress, Berlin 2006. Retrieved 2/22/2009 from http://www.ib1-unihh.de/IFSAM06.pdf.

Karande, K., and A. Merchant. The impact of time and planning orientation on an individual's recreational shopper identity and shopping behavior. (2012). *Journal of Marketing Theory and Practice,* 20(1) (Winter 2012) 59-72. DOI 10.2753/MTP1069-6679200104

Lennings, C. J., and A. M. Burns (1998). Time perspective: temporal extension, time extension and impulsivity. *Journal of Psychology,* 132 (July), 367–380.

Lusardi, A. (1999). Information, expectations and savings for retirement. *Behavioral Dimensions of Retirement Economics*. Brookings Institute and Russell Sage Foundation.

Murtagh, N., N. Lopes, and E. Lyons. (2012). Decision making in voluntary career change: An other-than-rational perspective. *The Career Development Quarterly*, 58(3): 249-263.

Nunally, J. C., and I. H. Bernstein. (1994). *Psychometric Theory*, (McGraw-Hill, New York), 3rd Edition.

Raju, P.S. (1980). Optimum stimulation level: its relationship to personality, demographics and exploratory behavior. *Journal of Consumer Research,* 7 (December), 272–282.

Riley, C. (2012). Family net worth plummets nearly 40%. *CNN Money.* http://money.cnn.com/2012/06/11/news/economy/fed-family-net-worth/index.htm

Sheppard, B. H., J. Hartwick, and P.R. Warsha.w (1988). The theory of reasoned action: a meta-analysis of past research with recommendations for modifications and future research, *Journal of Consumer Research*, 15 (December), 325-343.

Whitehouse, M. (2011). Number of the week: Companies cash hoard grows. *WSJ Blogs*. http://blogs.wsj.com/economics/2011/03/12/number-of-the-week-companies-cash-hoard-grows/

Wold, H. (1982). Soft modeling: the basic design and some extensions. In K. G. Joreskog and H. Wold (Eds.), *Systems under indirect observation: Causality, structure, prediction. 2,* 1-54. North-Holland: Amsterdam.

Zimbardo, P. G., and J. N. Boyd (1999). Putting time in perspective: a valid, reliable individual-differences metric. *Journal of Personality and Social Psychology,* 77 (6), 1271–1288.

Zimbardo, P., and J. Boyd. (2008). *The Time Paradox*, (Free Press, New York).

Appendix A, 2008/2009 Data

Today's date _____

Read each item and, as honestly as you can, answer the question: "How characteristic or true is this of me?" Circle the appropriate number, using the following scale:

1 = very untrue 2 = untrue 3 = neutral 4 = true 5 = very true DK = don't know

Please answer ***all*** the following questions.

1.	I believe that getting together with one's friends to party is one of life's important pleasures.	1 2 3 4 5 DK
2.	Familiar childhood sights, sounds, and smells often bring back a flood of wonderful memories.	1 2 3 4 5 DK
3.	Fate determines much in my life.	1 2 3 4 5 DK
4.	I often think of what I should have done differently in my life.	1 2 3 4 5 DK
5.	My decisions are mostly influenced by people and things around me.	1 2 3 4 5 DK
6.	I believe that a person's day should be planned ahead each morning.	1 2 3 4 5 DK
7.	It gives me pleasure to think about my past.	1 2 3 4 5 DK
8.	I do things impulsively.	1 2 3 4 5 DK
9.	If things don't get done on time, I don't worry about it.	1 2 3 4 5 DK
10.	When I want to achieve something, I set goals and consider specific means for reaching those goals.	1 2 3 4 5 DK
11.	On balance, there is much more good to recall than bad in my past.	1 2 3 4 5 DK
12.	When listening to my favorite music, I often lose all track of time.	1 2 3 4 5 DK
13.	Meeting tomorrow's deadlines and doing other necessary work come before tonight's play.	1 2 3 4 5 DK
14.	Since whatever will be will be, it doesn't really matter what I do.	1 2 3 4 5 DK
15.	I enjoy stories about how things used to be in the "good old times."	1 2 3 4 5 DK
16.	Painful past experiences keep being replayed in my mind.	1 2 3 4 5 DK
17.	I try to live my life as fully as possible, one day at a time.	1 2 3 4 5 DK
18.	It upsets me to be late for appointments.	1 2 3 4 5 DK
19.	Ideally, I would live each day as if it were my last.	1 2 3 4 5 DK
20.	Happy memories of good times spring readily to mind.	1 2 3 4 5 DK
21.	I meet my obligations to friends and authorities on time.	1 2 3 4 5 DK
22.	I've taken my share of abuse and rejection in the past.	1 2 3 4 5 DK
23.	I make decisions on the spur of the moment.	1 2 3 4 5 DK
24.	I take each day as it is rather than try to plan it out.	1 2 3 4 5 DK

25.	The past has too many unpleasant memories I prefer not to think about.	1 2 3 4 5 DK
26.	It is important to put excitement in my life.	1 2 3 4 5 DK
27.	I've made mistakes in the past that I wish I could undo.	1 2 3 4 5 DK
28.	I feel it's more important to enjoy what you're doing than to get work done on time.	1 2 3 4 5 DK
29.	I get nostalgic about my childhood.	1 2 3 4 5 DK
30.	Before making a decision, I weigh the costs against the benefits.	1 2 3 4 5 DK
31.	Taking risks keeps my life from becoming boring.	1 2 3 4 5 DK
32.	It's more important for me to enjoy life's journey than to focus only on the destination.	1 2 3 4 5 DK
33.	Things rarely work out as I expected.	1 2 3 4 5 DK
34.	It's hard for me to forget unpleasant images of my youth.	1 2 3 4 5 DK
35.	It takes joy out of the process and flow of my activities if I have to think about goals, outcomes, and products.	1 2 3 4 5 DK
36.	Even when I am enjoying the present, I am drawn back to comparisons with similar past experiences.	1 2 3 4 5 DK
37.	You can't really plan for the future because things change so much.	1 2 3 4 5 DK
38.	My life path is controlled by forces I cannot influence.	1 2 3 4 5 DK
39.	It doesn't make sense to worry about the future, since there is nothing that I can do about it anyway.	1 2 3 4 5 DK
40.	I complete projects on time by making steady progress.	1 2 3 4 5 DK
41.	I find myself tuning out when family members talk about the way things used to be.	1 2 3 4 5 DK
42.	I take risks to put excitement in my life.	1 2 3 4 5 DK
43.	I make lists of things to do.	1 2 3 4 5 DK
44.	I often follow my heart more than my head.	1 2 3 4 5 DK
45.	I am able to resist temptations when I know that there is work to be done.	1 2 3 4 5 DK
46.	I find myself getting swept up in the excitement of the moment.	1 2 3 4 5 DK
47.	Life today is too complicated; I would prefer the simpler life of the past.	1 2 3 4 5 DK
48.	I prefer friends who are spontaneous rather than predictable.	1 2 3 4 5 DK
49.	I like family rituals and traditions that are regularly repeated.	1 2 3 4 5 DK
50.	I think about the bad things that have happened to me in the past.	1 2 3 4 5 DK

51.	I keep working at difficult, uninteresting tasks if they will help me get ahead.	1 2 3 4 5 DK
52.	Spending what I earn on pleasures today is better than saving for tomorrow's security.	1 2 3 4 5 DK
53.	Often luck pays off better than hard work.	1 2 3 4 5 DK
54.	I think about the good things that I have missed out on in my life.	1 2 3 4 5 DK
55.	I like my close relationships to be passionate.	1 2 3 4 5 DK
56.	There will always be time to catch up on my work.	1 2 3 4 5 DK
57.	In **Fall 2007** I felt that my savings in a bank were secure.	1 2 3 4 5 DK
58.	In **Fall 2007** I felt that my investment funds (stocks & bonds) were secure.	1 2 3 4 5 DK
59.	In **Fall 2007** I felt that my job (source of income) was secure.	1 2 3 4 5 DK
60.	In **August 2008** I felt that my savings in a bank were secure	1 2 3 4 5 DK
61.	In **August 2008** I felt that my investment funds (stocks & bonds) were secure.	1 2 3 4 5 DK
62.	In **August 2008** I felt that my job (source of income) was secure.	1 2 3 4 5 DK
63.	**Today** I feel that my savings in a bank is secure.	1 2 3 4 5 DK
64.	**Today** I feel that my investment funds (stocks & bonds) are secure.	1 2 3 4 5 DK
65.	**Today** I feel that my job (source of income) was secure.	1 2 3 4 5 DK
66.	As a result of changes in the economy many of my relatives are moving their financial assets from financial markets into banks.	1 2 3 4 5 DK
67.	As a result of changes in the economy many of my relatives are moving their financial assets from financial assets into cash.	1 2 3 4 5 DK
68.	As a result of changes in the economy many of my relatives are looking for a new job	1 2 3 4 5 DK
69.	As a result of changes in the economy many of my relatives are retiring.	1 2 3 4 5 DK
70.	As a result of changes in the economy many of my relatives are training for a new job.	1 2 3 4 5 DK
71.	As a result of changes in the economy many of my relatives are - _____ (please specify and state extent to which it true.	1 2 3 4 5 DK
72.	As a result of how I feel now, I intend to move my financial assets from financial markets into banks.	1 2 3 4 5 DK

73.	As a result of how I feel now, I intend to move my financial assets from financial assets into cash.	1 2 3 4 5 DK
74.	As a result of how I feel now, I intend to look for a new job.	1 2 3 4 5 DK
75.	As a result of how I feel now, I intend to retire.	1 2 3 4 5 DK
76.	As a result of how I feel now, I intend to train for a new job.	1 2 3 4 5 DK
77.	As a result of how I feel now, I intend to _____ (Please specify and state extent to which it is true.)	1 2 3 4 5 DK
78.	I have the power to improve my current financial situation.	1 2 3 4 5 DK
79.	I understand what is going on in the economy.	1 2 3 4 5 DK
80.	I understand what is going on in the financial markets.	1 2 3 4 5 DK

DEMOGRAPHICS: Circle the number that corresponds to the category that best describes you:

Sex: 1. Male 2. Female **Age at last birthday_____ Zip code_____**

I own my own business. 1. Yes 2. No **I'm a partner in a business**. 1. Yes 2. No

I do independent consulting work. 1. Yes 2. No

I work in _____industry

Currently taking college classes? 1. Yes 2. No

Your major (college students only) _____

Current household monthly income (approximately) _____

Approximate dollar value of your financial assets (savings, investments etc.)?

1. less than 25,000
2. 25,000-49,999
3. 50,000- 74,999
4. 75,000- 99,999
5. 100,000- 124,999
6. 125,000-149,999
7. 150,000-174,999
8. 175,000-199,999
9. 200,000-224,999
10. 225,000- 249,999
11. 250,000 – 274,999
12. 275,000 – 299,999
13. 300,000 – 324,999
14. 325,000 – 349,999
15. 350,000 – 374,999
16. 375,000 – 399,999
17. 400,000 – 424,999
18. 425,000 – 449,999
19. 450,000 – 474,999
20. 475,000 – 499,999
21. 500,000 +

IN PERCENTAGES, how your <u>financial assets</u> are distributed among the following? (must add up to 100%)

1. Checking accounts _____ 2. Savings accounts _____

3. Stocks/bonds/mutual funds _____ 4. Retirement/pension funds_____

5. Other _____ (please specify _____)

IN PERCENTAGES, how your <u>real assets</u> are distributed among the following? (must add up to 100%)

1. Home_____ 2. Vehicles _____3. Other real estate _____

4. Personal property (furniture, tools electronics, jewelry, etc.)_____

5. Other _____ (please specify: _____)

Highest level of educational attainment: 1. Less than high school 2. High school/GED

3. Some college 4. Bachelors degree 5. Masters degree 6. Above Masters degree

What is the subject area is your highest degree (college graduates only)?

I would classify my business experience level as:

1. Very Low 2. Low 3. Average 4. High 5. Very High

What is your Racial/ethnic identity?

1. African American 2. Asian American 3. Hispanic American

4. Native American 5. White American

What is your religious affiliation?

1. Catholic 2. Protestants (all Christian denominations that are not Catholic)

3. Jewish 4. Muslim 5. Atheist

6. Other (please specify)_____

How many times do you pray (on your own) weekly? _____

How many times do you attend a religious activity (church etc)? _____

How important is religion in your personal decisions?

1. Very unimportant 2. Unimportant 3. Important 4. Very Important

How would you describe yourself politically?

Very Liberal			Moderate			Very Conservative
1	2	3	4	5	6	7

Appendix B, Results for ZTPI with the 2008/2009 Data

Approximately 458 members of a South Texas university community (including students, faculty members, and administrators/staff) participated in this study, producing 451 usable responses for the 2008/2009 Model. Data were collected through surveys that included only a few open-ended questions. Respondents from each of the categories (students, faculty members and administrators/staff) were selected both purposefully and on the basis of convenience. This means, for example, professors teaching classes of at least 60 students were more likely to be solicited for permission to administer the surveys in their classes than those with smaller classes.

Care was taken, nevertheless, to ensure that participating students were distributed from freshmen through graduates. Similar care was exercised to ensure that every college in the university was represented, in terms faculty and staff participation. The university's administrative offices were specifically targeted to ensure the inclusion of senior administrative personnel in the sample. Unlike those distributed to students, and faculty, staff and administrators' questionnaires were mailed with two separate return, self-addressed envelopes for separately returning the questionnaires and Informed Consent Forms. Overall, the sample reflects the general demographic distribution of the university.

Results Summary of the 2008/2009 ZTPI Model

The median age of the participants was 23 years, with a range from 16 to 71. Fifty-nine percent of the participants were female. Respondents' experience varied from very little perceived business experience to extensive business experience, with the average participant's self-rating as an experience level of 3.0 on a 5-point scale, similar to national averages. Household income was at least $2,000 per month for approximately 71 percent of the respondents, and the average monthly income was $4,818.

2008/2009 Model

The results of the confirmatory factor analysis suggest that the measurement items within each scale are only moderately correlated with the underlying latent variable. Less than half of the average variance for each factor is explained, with the exception of the two intent items. This indicates that the measurement items in these scales exhibits only moderate convergent validity and are highly correlated to each other due to a single underlying construct. "In an adequate model for exploratory purposes, composite reliabilities should be greater than 0.6" (Chin, 1998; Höck and Ringle, 2006). The average variance explained is summarized in Table B2, along with the results of composite reliability tests for each of the scales. Except for the Future construct, each of the reliability statistics generally approaches or exceeds the 0.60 recommended by Chin (1998) and Höck and Ringle (2006). Table B3 shows the correlations among the latent variables; the numbers presented in the diagonal depict the square root of the average common variance extracted by the measurement items within the scale (the average inter-item correlation). Based on the results, the measurements exhibit moderate validity and reliability. The path coefficients to the dependent variables from the latent variables are presented in Figure 1B with commentary.

Three path coefficients are significant at $p \leq 0.05$, and of the expected signs for supporting ZTPI construct hypotheses: A strong past-positive to Intent to Move Job is negative.

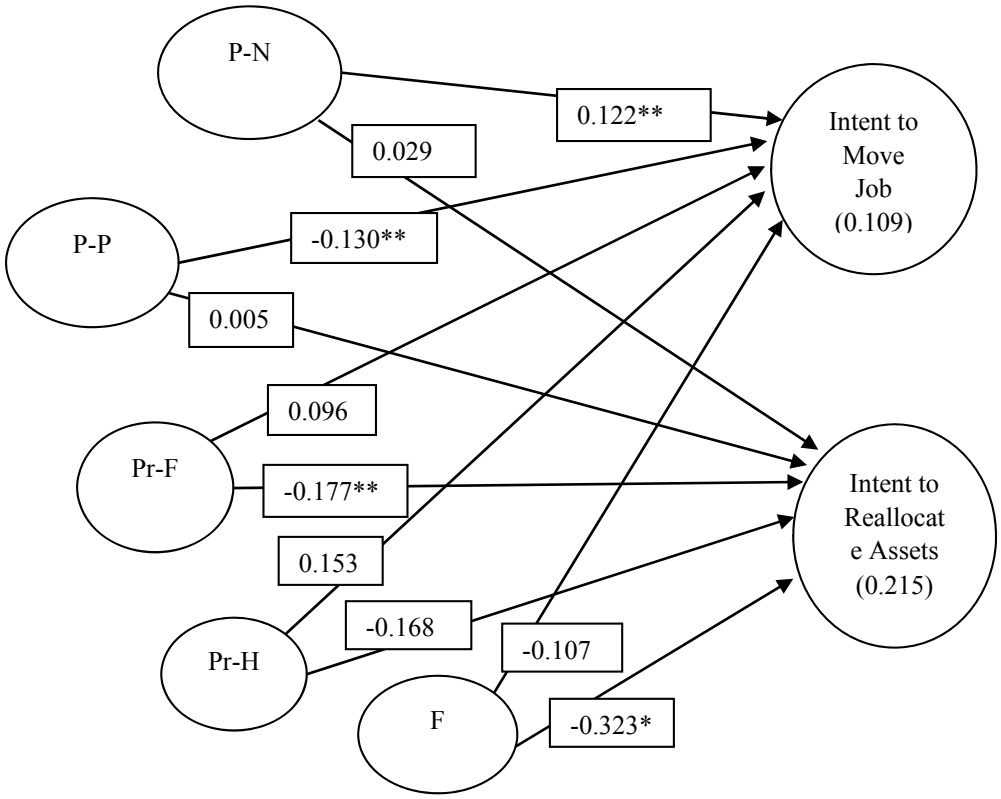
- A strong past-negative to Intent to Move Job is positive.
- A strong present-fatalistic to Intent to Reallocate Assets is negative.

The negative effect of strong Present-Fatalistic orientation is significant; however a positive sign, not a negative one was predicted. See Figure 1B. Of the two time perspectives for which no effect was expected, that was indeed the outcome:

- A strong past-negative orientation had *no* effect on intent to move funds.
- A strong present-fatalistic orientation had *no* effect on the intent to change jobs.

The path, representing Future orientation to the intent to reallocate assets, approached significance, with $p \leq 0.0635$, and it had the expected, negative sign. All other paths, representing hypotheses, were insignificant.

Figure B1 – 2008/2009 Model Structural Model Using Partial Least Squares



* Marginally significant at p = 0.0635
** Significant at $p \leq 0.050$.

An overall absolute value of a path coefficient for the dependent variable that is greater than 0.20 is considered meaningful (Chin, 1998). One path, from Future to Intent to Reallocate Assets, was greater than 0.20, and two paths approached this size: the path from Present Fatalistic to Intent to Reallocate Assets and the path from Present Hedonistic to Intent to Reallocate Assets. Only (1) the pathway from Past Negative orientation and (2) and pathway from Past Positive orientation were significant, with $p < 0.05$.

The amount of variance in the dependent variables explained by Model 1 is represented by the squared multiple correlations of 0.109 for intent to change jobs and 0.215 for intent to reallocate assets. That is, taken as a set, these constructs explained about one-tenth and one-fifth of the total

behavior, respectively. Chin (1998) and Höck and Ringle (2004) consider a cut-off of 0.33 to be moderate evidence, and a cut-off of 0.19 to be weak substantiation, respectively.

Of the control variables, the path from age to moving jobs was significant and negative, but the path size was -0.189, indicating that older individuals were less likely to change jobs. This could reflect loyalty to their current company, or a perceived disadvantage (age discrimination) in the work place.

Table B1. 2008/2009 Model, ZTPI Variables Using Primary Least Squares

Factor	Item #	Factor Loading	Weight	Factor	Item #	Factor Loading	Weight
Intent-Reallocate Assets	**71**	**0.8264**	**0.5095**	**Past Positive**	**2**	**0.5970**	**0.2502**
Intent-Reallocate Assets	**72**	**0.9085**	**0.6788**	**Past Positive**	**7**	**0.7389**	**0.2819**
Intent - Move Job	73	0.9003	0.7158	**Past Positive**	**11**	**0.7619**	**0.3859**
Intent – Move Job	74	0.3577	0.1933	**Past Positive**	**15**	**0.3788**	**0.0851**
Intent – Move Job	75	0.7910	0.3967	**Past Positive**	**20**	**0.7359**	**0.2915**
Present Hedonistic	**1**	**0.2867**	**0.0820**	**Past Positive**	**29**	**0.2308**	**0.0039**
Present Hedonistic	**8**	**0.4690**	**0.1282**	**Past Positive**	**49**	**0.4950**	**0.2090**
Present Hedonistic	**9**	**0.0187**	**-0.0380**	Past Negative	4	0.5225	0.0530
Present Hedonistic	**12**	**0.2415**	**0.0118**	Past Negative	5	0.3110	0.0937
Present Hedonistic	**17**	**0.5906**	**0.2071**	Past Negative	16	0.7764	0.1716
Present Hedonistic	**19**	**0.4944**	**0.0987**	Past Negative	22	0.5720	0.1748
Present Hedonistic	**23**	**0.4295**	**0.0428**	Past Negative	25	0.7114	0.2181
Present Hedonistic	**24**	**0.2402**	**-0.0076**	Past Negative	27	0.4929	0.1157
Present Hedonistic	**26**	**0.5952**	**0.1387**	Past Negative	33	0.4785	0.1184
Present Hedonistic	**28**	**0.1667**	**-0.0670**	Past Negative	34	0.7598	0.2239
Present Hedonistic	**31**	**0.6652**	**0.1618**	Past Negative	36	0.4447	0.0427
Present Hedonistic	**32**	**0.5223**	**0.1488**	Past Negative	50	0.7882	0.1744
Present Hedonistic	**42**	**0.6991**	**0.1798**	Past Negative	54	0.6859	0.1728
Present Hedonistic	**44**	**0.4876**	**0.2110**	**Future**	**6**	**0.0161**	**-0.1054**
Present Hedonistic	**46**	**0.6819**	**0.1889**	**Future**	**10**	**0.3588**	**0.1802**
Present Hedonistic	**48**	**0.5172**	**0.1691**	**Future**	**13**	**0.2264**	**0.1224**
Present Hedonistic	**55**	**0.3864**	**0.0837**	**Future**	**18**	**0.5456**	**0.3379**
Present Hedonistic	**56**	**0.0043**	**-0.1551**	**Future**	**21**	**0.5756**	**0.4519**
Present Fatalistic	3	0.3385	0.1148	**Future**	**30**	**0.5113**	**0.3710**
Present Fatalistic	14	0.5837	0.1451	**Future**	**40**	**-0.2131**	**-0.3926**
Present Fatalistic	35	0.4643	0.1404	**Future**	**43**	**0.0832**	**-0.0755**
Present Fatalistic	37	0.5980	0.2345	**Future**	**45**	**0.1260**	**0.0760**
Present Fatalistic	38	0.6280	0.2730	**Future**	**51**	**-0.3918**	**-0.4852**
Present Fatalistic	39	0.6752	0.2968				
Present Fatalistic	41	0.4277	0.2191				
Present Fatalistic	47	0.1576	0.0360				
Present Fatalistic	52	0.4724	0.1329				
Present Fatalistic	53	0.5737	0.2429				

Table B2. 2008/2009 Model Common Variance Explained and Composite Reliability Measures

Construct	Average Variance Explained	Composite Reliability
Intent – Reallocate Assets	0.754	0.860
Intent - Move Job	0.521	0.745
Past Positive	0.352	0.774
Past Negative	0.377	0.862
Present Hedonistic	0.217	0.799
Present Fatalistic	0.264	0.767
Future	0.129	0.516

Table B3. 2008/2009 Model Correlations among Latent Variables

Construct	Intend to Reallocate Assets	Intend to Move Job	Past Positive	Past Negative	Present Positive	Present Negative	Future
Intent – Reallocate Assets	**0.868**						
Intent - Move Job	0.233	**0.722**					
Past Positive	-0.189	-0.183	**0.593**				
Past Negative	0.143	0.233	-0.314	**0.614**			
Present Hedonistic	-0.210	0.118	0.267	0.080	**0.466**		
Present Fatalistic	0.253	0.225	-0.244	0.409	0.126	**0.514**	
Future	-0.409	-0.161	0.300	-0.176	0.209	-0.268	**0.359**

* The numbers presented in the diagonal depicting the square root of the average common variance extracted by the measurement items within the scale.

This model was more effective in predicting whether one would move money or reallocate assets than whether one would change jobs. This could be in part because one has more control over moving money; a job change requires an accommodating job market. Additionally, people did not admit to relying on past experience or teaching to adapt to the (then) present economic crisis. Their future beliefs were sizably influential, but less stable of a predictor than present indicators. All indicators were from the standard ZTPI. Many of these items are very general in nature. Had, for example, the item "I meet my obligations to friends and authorities on time" been tailored to read "I meet my financial obligations on time," a tighter, more consistent result might have been obtained.

Appendix C, Twenty-one Additional TPIF Questions for the 2010, Intent to Change Jobs Model

Read each item and, as honestly as you can, answer the question: "How characteristic or true is this of me?" Circle the appropriate number, using the following scale:

1 = very untrue 2 = untrue 3 = neutral 4 = true 5 = very true

Please answer ***all*** the following questions.

81.	I believe that spending money to party with one's friends is one of life's important pleasures.	1 2 3 4 5
82.	I believe that a person's budget should be planned ahead each morning.	1 2 3 4 5
83.	If things don't get paid on time, I don't worry about it.	1 2 3 4 5
84.	It gives me pleasure to think about my savings.	1 2 3 4 5
85.	When I want to have money, I set goals and consider specific means for reaching those goals.	1 2 3 4 5
86.	Meeting tomorrow's payments and doing other necessary work comes before tonight's play.	1 2 3 4 5
87.	I believe that my future finances are well planned.	1 2 3 4 5
88.	I spend and earn in order to live my life as fully as possible one day at a time.	1 2 3 4 5
89.	It doesn't make sense to worry about future income security since there is nothing to do about it anyway.	1 2 3 4 5
90.	When I have money, I like playing and betting.	1 2 3 4 5
91.	It upsets me to be late in making payments.	1 2 3 4 5
92.	I spend impulsively and I make purchasing decisions at the moment.	1 2 3 4 5
93.	I feel that it's more important to enjoy what you're doing than to save on a schedule.	1 2 3 4 5
94.	I'm inclined to spend impulsively when feeling emotional.	1 2 3 4 5
95.	It upsets me when people who owe me money are late paying me back.	1 2 3 4 5
96.	When I go shopping I lose control over how much I spend.	1 2 3 4 5
97.	I save on schedule by making steady progress.	1 2 3 4 5
98.	I gamble to put excitement in my life.	1 2 3 4 5
99.	I write down a budget to plan my finances.	1 2 3 4 5

100. I keep working at difficult, uninteresting tasks if they will help me get ahead financially.	1 2 3 4 5
101. I am able to resist temptations when I know that I need to save.	1 2 3 4 5

Appendix D, 2012 TPIF2 Questions for Intent to Change Jobs Model

1. Spending money to be with one's friends is one of life's important pleasures.	1 2 3 4 5 DK
2. Memorabilia from my childhood often brings back many wonderful memories.	1 2 3 4 5 DK
3. My income and wealth determine much of my life.	1 2 3 4 5 DK
4. I often think of how I should have invested or spent my money differently.	1 2 3 4 5 DK
5. My purchases and savings are mostly influenced by people and things around me.	1 2 3 4 5 DK
6. A person's spending should be planned ahead.	1 2 3 4 5 DK
7. It gives me pleasure to think about things I've bought.	1 2 3 4 5 DK
8. I buy things impulsively.	1 2 3 4 5 DK
9. If I don't make payments on time, I don't worry about it.	1 2 3 4 5 DK
10. When I want to purchase something, I set goals and specific means to reach those goals.	1 2 3 4 5 DK
11. Overall, there is much more good than bad to recall in my past spending and/or saving.	1 2 3 4 5 DK
12. When using my favorite purchases, I often lose track of time.	1 2 3 4 5 DK
13. Meeting payment deadlines and other necessary expenses come before tonight's play.	1 2 3 4 5 DK
14. Since whatever will be will be, it doesn't really matter how I spend my money.	1 2 3 4 5 DK
15. I enjoy stories about how inexpensive things used to be in the "good old times."	1 2 3 4 5 DK
16. Regrets about past purchases and investments keep being replayed in my mind.	1 2 3 4 5 DK
17. It upsets me to be late in making payments.	1 2 3 4 5 DK
18. Ideally, I would spend money each day as if it were my last.	1 2 3 4 5 DK
19. Happy memories of making good deals on my purchases spring readily to mind.	1 2 3 4 5 DK
20. I meet my obligations and make my payments on time.	1 2 3 4 5 DK

21. I've had my share of bad jobs and unemployment in the past.	1 2 3 4 5 DK
22. I make spending decisions on the spur of the moment.	1 2 3 4 5 DK
23. I spend my money as it comes in rather than trying to plan my spending.	1 2 3 4 5 DK
24. The past has too many unpleasant memories of not having enough money that I prefer not to think about.	1 2 3 4 5 DK
25. It is important to spend money to add excitement to my life.	1 2 3 4 5 DK
26. I've made mistakes in the past with my money that I wish I could undo.	1 2 3 4 5 DK
27. I feel it's more important to enjoy your money today rather than save it for later.	1 2 3 4 5 DK
28. I get nostalgic about all the things I had during my childhood.	1 2 3 4 5 DK
29. Before making a purchase, I weigh the costs against the benefits.	1 2 3 4 5 DK
30. Taking risks with my money keeps my life from becoming boring.	1 2 3 4 5 DK
31. It's more important for me to enjoy spending than to focus on only what I am accumulating.	1 2 3 4 5 DK
32. I rarely stay within my budget.	1 2 3 4 5 DK
33. It's hard for me to forget unpleasant images related to money issues from my youth.	1 2 3 4 5 DK
34. It takes joy out of having money if I think about finances, savings, and goals.	1 2 3 4 5 DK
35. When I spend money, I am drawn back to comparisons with similar past purchases.	1 2 3 4 5 DK
36. You can't really plan for future expenses and needs because things change so much.	1 2 3 4 5 DK
37. My income is controlled by forces I cannot influence.	1 2 3 4 5 DK
38. It doesn't make sense to worry about having enough money in the future, since there is nothing I can do about it anyway.	1 2 3 4 5 DK
39. I reach my financial goals on schedule by making steady progress.	1 2 3 4 5 DK
40. I find myself tuning out when family members talk about the way their finances used to be.	1 2 3 4 5 DK
41. I make risky investments to put excitement in my life.	1 2 3 4 5 DK
42. I make budgets and lists of my anticipated expenses.	1 2 3 4 5 DK
43. I often spend money based on my heart more than my head.	1 2 3 4 5 DK

44.	I am able to resist temptations when I know there is a need to save money.	1 2 3 4 5 DK
45.	When I spend money, I find myself getting swept up in the excitement of the moment.	1 2 3 4 5 DK
46.	Finances today are too complicated; I would prefer the simpler choices of the past.	1 2 3 4 5 DK
47.	I prefer friends who spend spontaneously rather than predictably.	1 2 3 4 5 DK
48.	I like to spend money to continue family traditions.	1 2 3 4 5 DK
49.	I think about the bad financial circumstances that have hurt me in the past.	1 2 3 4 5 DK
50.	I stick to a budget if it will help me get ahead.	1 2 3 4 5 DK
51.	Spending what I earn on pleasures today is better than saving for tomorrow's security.	1 2 3 4 5 DK
52.	Often luck pays off better than hard work.	1 2 3 4 5 DK
53.	I think about the financial opportunities I have missed out on in my life.	1 2 3 4 5 DK
54.	I like to surprise those close to me with gifts.	1 2 3 4 5 DK
55.	There will always be time to invest and save money.	1 2 3 4 5 DK
56.	As a result of how I feel now, I intend to move my financial assets from financial markets into banks.	1 2 3 4 5 DK
57.	As a result of how I feel now, I intend to move my financial assets from financial assets into cash.	1 2 3 4 5 DK
58.	As a result of how I feel now, I intend to look for a new job.	1 2 3 4 5 DK
59.	As a result of how I feel now, I intend to retire.	1 2 3 4 5 DK
60.	As a result of how I feel now, I intend to train for a new job.	1 2 3 4 5 DK