Volume 15, Number 2

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

IMPACTS OF JA BIZTOWN ON IMPROVING FINANCIAL LITERACY AMONG MIDDLE-SCHOOL STUDENTS

Jack G. Brancewicz, Junior Achievement of Greater New Orleans, Inc.
Juli Pattison, Junior Achievement of Greater New Orleans
Lillian Y. Fok, University of New Orleans

ABSTRACT

There were 1329 students from 22 schools participated in the JA BizTown program and 90 students from two schools were in the control group. Comparisons of the JA BizTown and control group were made in terms of increase in matched paired objective test (30 questions), increase in matched paired students' attitude, overall pre-test and post-test scores, and overall pre-test and post-test attitudinal scores. The MANOVA results indicated that JA BizTown curriculum can dramatically increase the students' knowledge in Finance and Economics when comparing the BizTown group and the control group. Furthermore, JA BizTown curriculum can dramatically increase the students' knowledge in Finance and Economics regardless of their gender and ethnicity.

INTRODUCTION

Low financial literacy among American students causes serious concern among educators and education advocacy groups [ACEC, 2002; ASEC, 1999]. Because of the current economic situation, more schools realize the importance and urgency for students to acquire literacy in economics and finance during their earlier years [American School Board Journal, 2008; Black, 2009; Finkel, 2010; Varcoe et al., 2005]. The Oklahoma legislature signed The Passport to Financial Literacy Act in 2007, requiring all seventh-graders demonstrate proficiency in 14 financial areas [Black, 2009]. Different states around the country and the companies in the private sector have developed similar finance programs for schools.

The program used in this study is JA BizTown. JA BizTown is designed to inspire and prepare 5th and 6th grade students for a lifetime of learning and academic achievement through career exploration and financial literacy. Research shows that career development begins in early childhood and peaks at age 10 when students often model their behavior and career aspirations after their parents [Auger, 2005]. According to a 2009 report from a U.S. Census survey in the metro area, 47% of area workers earn less than a sustainable, self-sufficiency wage of $35,000 annually for a family of four, and Louisiana is ranked 2nd in the nation where a child is more likely to become homeless. In Orleans Parish, an average of 81% of our students is

receiving free or reduced lunch benefits as a result of poverty. Through JA BizTown for 5th and 6th grade students, Junior Achievement (JA) provides memorable, motivational experiences to this vulnerable population that allow them to look beyond their current circumstances and envision a life of self-sufficiency, in which needs are met and dreams become reality. JA BizTown strengthens student assurance in their own abilities to approach challenges and master challenges with success. With so many or our children in metro area schools from disadvantaged families, a child's self-efficacy -- his belief in himself to achieve success, socially and economically in adult life -- becomes key to his success – as a student, as a consumer and as a future leader in our society. These motivational factors become "rooted in the core belief that one has the power to produce effects by one's actions" [Bandura et al, 2001]. Through hands-on, in-class learning in daily lessons, a simulation experience where students construct knowledge and value from direct experiences, and reflection on personal actions, students develop an understanding of the relationship between what they learn in school and confidence in their own ability to succeed in the classroom and beyond.

In case studies of adults in financial crisis, financial literacy education is vital to their successful recovery. In New Orleans, the United Way, Project Reach, the Salvation Army and many other organizations provide adult financial literacy education in recognition of its importance. JA is confident that financial literacy education ***before*** a crisis occurs will help alleviate this demand on our social structure and inspire our children to go beyond their disadvantage circumstances toward a brighter future. JA BizTown was piloted in nine cities, including New Orleans, and its proven curriculum equips our young people with the skills they need to achieve this future. By providing the simulation experience, JA creates an experiential learning environment that has a connection to the real world beyond the school walls and provides a memorable experience in financial literacy – a memory that will last a lifetime through their future decisions as workers, community leaders and consumers. Our children study art; however, art becomes a living experience when a paint brush and art supplies are given to the child. The experience of viewing a painting in a museum is just as memorable. It is only through JA BizTown that our children will experience the business of life – working as an employee or supervisor, balancing a checkbook, paying loans, being responsible for their safety and the safety of others and much more in the real world of work. The JA curriculum is based on tested best practices and proven educational theory and has the buy-in of stakeholders and school leaders. The educator-led JA BizTown curriculum and implementation strategies were designed to be easily managed to ensure faithful treatment of the educator mandates required of the program. JA has developed strategies, based on tested best practices, to strengthen in-class learning and site simulation experiences. While all JA K-12 programs directly reflect core curricula and state mandated *Grade Level Expectations* in English, language, mathematics and social studies, the JA BizTown curriculum has been intensely researched to ensure direct correlation because of its extended use of crucial in-class time. Experiential learning recognizes that optimal learning is achieved through moving beyond knowledge and learning a skill to actually using one's knowledge and skill in a practical experience. The curriculum was

developed by JA Worldwide educational staff using evaluative data from field tests and proven themes of contemporary education with focus on the needs of children, educators and communities: experiential learning, constructivism and collaborative learning. The national field tests showed students outperformed the comparison group in content knowledge, were more likely to improve grade scores, and were highly engaged in learning activities. Educators reported the lessons were effective in accommodating different learning styles.

To ensure the continuity of the program, all curricular materials target the seven goals below. Lesson objectives are directly related to the goals, and assessments are directly tied to the lesson objectives. In addition, JA BizTown goals and objectives are consistent with the appropriate state GLEs for 5th and 6th grades, and the lessons reflect similar instructional approaches to those found in the Louisiana Comprehensive Curriculum. Several indirect effects are also anticipated, including but not limited to building essential life skills, improving critical thinking and reasoning skills, and developing knowledge in core curricular areas such as English and language arts, math, and social studies.

Goal 1: Students can explain the roles citizens play in their community as workers and consumers and relate these roles to the free enterprise system. Students will synthesize information from these lessons and in the simulation. Assessment tools: tests from Units 1, 3, 4, and 5 and the post-test.

Goal 2: Students can explain the importance of citizens' rights and responsibilities in a community. Students will reason based on their ability to generalize from evidence accumulated in the lessons and the simulation. Assessment tools: tests from Units 1, 3, and 4 and the post-test.

Goal 3: Students can demonstrate a basic understanding of the free enterprise system. This goal is based on the students' knowledge of the free enterprise system as presented across all units and focused on their performance during the simulation. Assessment tools: unit tests, post-test, and observations of the students' performance during their visit to JA BizTown using a protocol developed by the evaluator.

Goal 4: Students can explain the importance of non-profit organizations in our communities. This goal is addressed in Unit 1 and requires students to synthesize their knowledge of non-profit organizations with the roles these organizations play in a free enterprise system. Assessment tools: Unit 1 test and post-test.

Goal 5: Students can demonstrate money management skills. This goal is addressed by three units. It requires students to know specific banking practices (e.g., opening a bank account, depositing money, writing checks, maintaining balances, etc.), as well as apply this knowledge. Assessment tools: tests from Units 2, 3, and 4, the post-test, and data collected during the simulations.

Goal 6: Students can demonstrate basic business practices and responsibilities. This goal is the focus of Unit 4. It requires students to know specific information related to business

management and apply this knowledge. Assessment tools: test for Unit 4, post-test, and observations of the students during their visit to JA BizTown using a protocol developed by the evaluator.

Goal 7: Students can identify the soft skills necessary for successful participation in the world of work as well as demonstrate their use of them. This goal is addressed in Units 3 and 5 and requires students to know and use specific interpersonal skills during the simulation. Assessment tools: tests from Units 3 and 5, post-test, and observations of students' at the simulation using a protocol developed by the evaluator.

RESEARCH MODEL AND HYPOTHESES

Many researchers have studied and documented the financial literacy of students but few have actually evaluated the effectiveness of the programs empirically. Our study will adopt the model proposed by Becker and Walstad [1987]. The model proposes that cognitive achievement, gain in economics and finance knowledge, is affected by students' demographic characteristics and classroom/environmental influences. Student demographic characteristics will include gender, racial affiliation, family income (participation in subsidized lunch program), previous participation in other Junior Achievement programs, and attitude towards education. Classroom/environmental influences will include the number of teachers involved in delivering the JA BizTown curriculum, perception of time spent on lessons, and perception of teacher effectiveness.

The preceding discussion forms the underlying logic of the research question and hypotheses. In this study, the research question is to find out what are the important factors that affect students' economic understanding. Specifically, the current study will test the following hypotheses (**stated in alternative hypothesis**).

H_1 *Different gender groups have different levels of students' understanding of economics and financial concepts.*

H_2 *Different racial groups have different levels of students' understanding of economics and financial concepts.*

H_3 *Previous Junior Achievement participation is related to students' understanding of economics and financial concepts.*

H_4 *Students' attitude towards education is related to their understanding of economics and financial concepts.*

H_5 *The number of teachers involved in teaching the JA BizTown lessons is related to students' understanding of economics and financial concepts.*

H_6 Students' perception of time spent on lessons is related to their understanding of economics and financial concepts.

H_7 Students' perception of teacher effectiveness is related to their understanding of economics and financial concepts.

METHODOLOGY AND ANALYSIS

Research Design

This study is quasi-experimental, with pre-post treatment (JA BizTown) design. The sampling of schools into the experimental and control groups will not be randomized due to the intense participation required by the schools. The experimental group is comprised of the schools willing to participate in the JA BizTown program and the control group is comprised of schools with similar student profile but not participate in JA BizTown program. The target population includes 5^{th}-6^{th} grade students enrolled in public and charter schools in Orleans, Jefferson, Plaquemines, St. Bernard, and St. Tammany parishes during the academic year of 2011-2012.

Treatment/Intervention – BizTown

The JA Capstone Education Manager will direct and oversee the program, ensure outcome measures are in place and outcomes are recorded, compile quarterly reports on the progress of the program, and supervise program personnel. Program components include the following stages: Planning, Training, Classroom Education, Simulation and Debriefing (Appendix 1). The students will be tested for their economic and financial literacy and the same instrument will be administered again after two months.

Subjects

To examine the factors influencing the students' understanding of economics and finance concepts, all students in the (5^{th} and/or 6^{th} grade) class chosen by the participating school were tested before and after the JA BizTown program. There were 1329 students from 22 schools in the JA BizTown group and 90 students from two schools in the control group. In the JA BizTown group, there were 609 students from Orleans Parish, 359 from Jefferson, 262 from St. Tammany, and 99 from Plaquemines with 49% female (502) and 51% male (520). With respect to ethnicity, 41% indicated White or part White, 50% Black, 10% Hispanic, 6% Asian, 11% American Indian/Alaskan Native, and 12% Others. In the control group, there were 38 students

from Orleans Parish and 52 from Jefferson, with 45% female (28) and 55% male (34). With respect to ethnicity, 45% indicated White or part White, 40% Black, 24% Hispanic, 7% Asian, 10% American Indian/Alaskan Native, and 18% Others. The findings have indicated that the demographic background of the JA BizTown group and the control group are quite similar.

Instruments and Measurement of Variables

Economics and Financial Literacy Survey: It contains 30 multiple choice items developed based on important economic and financial concepts. This instrument was tested and validated in previous JA programs. This instrument will be used for pre-test and post-test. See Figure 1 for the summary of instruments and research variables.

Student Pre-Program and Post-Program Surveys: These surveys were created by the evaluator according to the established procedures for developing sound instrument. The Pre-Program Survey contains demographic questions and 11 perceptual questions on students' attitude and self-image using a 4-point Likert scale with "1" being "Strongly Disagree" and "4" being "Strongly Agree" (Figure 1). It will be given to the students when the pretest is administered. The Post-Program Survey contains the 11 perceptual questions on students' attitude and self-image in the Pre-Program Survey along with eight perceptual questions on JA BizTown curriculum design and delivery and seven questions on JA BizTown visit. The perceptual questions all use a 4-point Likert scale with "1" being "Strongly Disagree" and "4" being "Strongly Agree."

Figure 1: Summary of Instruments and Research Variables

Subject	Instrument	Variables	When
Student	Economics and Financial Literacy Survey	Understanding of Economic and Financial Concepts	Pre- and post-JA BizTown
Student	Student Pre-Program Survey	JA Participation Ethnicity Attitude towards education Self-evaluation of skills	Pre-JA BizTown
Student	Student Post-Program Survey	Gender Attitude towards education Self-evaluation of skills Perception on curriculum design and delivery Perception on BizTown visit	Post-JA BizTown

RESULTS AND CONCLUSIONS

Comparisons of the JA BizTown and control group were made in terms of increase in matched paired objective test (30 questions), increase in matched paired students' attitude,

overall pre-test and post-test scores, and overall pre-test and post-test attitudinal scores. Increase in objective test score and increase in student attitude score were computed by subtracting the pre-test score from the post-test score of the same student (matched-pair). MANOVA was used to determine if there are any differences between the two groups (JA BizTown vs. control).

The MANOVA results have indicated statistically significant higher increase in the objective test scores, 4.69 for JA BizTown group vs. 0.87 for the Control Group but no significant increase in students' attitude (Fiture 2). The results indicated that JA BizTown curriculum can dramatically increase the students' knowledge in Finance and Economics. Changing students' general attitude towards learning and their confidence levels may need a program that lasts, perhaps, not just a semester in order to bring deep rooted attitudinal changes. We also found significantly higher post-test scores (not matched) in JA BizTown group than the Control Group which is not surprising because JA BizTown is designed to improve the students' financial and economics knowledge. Since the JA BizTown group scored differently from the control group, the remaining analyses will focus on the JA BizTown group only.

Figure 2. Summary of MANOVA results

Treatment (JA BizTown n = 1329 and Control Group n = 90)

Dependent Variable	Treatment	Mean
Pre-test attitude	JA BizTown	3.384
	Control Group	3.372
Post-test attitude	JA BizTown	3.095
	Control Group	3.258
Pre-test scores	JA BizTown	11.577
	Control Group	10.906
Post-test scores	JA BizTown	16.270
	Control Group	11.774
Increase in objective test scores	JA BizTown	4.693
	Control Group	.868
Increase in student attitude scores	JA BizTown	-.288
	Control Group	-.115

Analysis of gender

Hypothesis 1 suggested that there would be no difference between male and female students in their understanding of and attitude towards economics and financial concepts. Male and female students were compared in terms of increase in matched paired objective test (30 questions), increase in matched paired students' attitude, overall pre-test and post-test scores, and overall pre-test and post-test attitudinal scores using MANOVA.

There are 502 female and 509 male students in the JA BizTown group who completed the JA BizTown curriculum. The MANOVA results indicated there is not enough evidence to find statistically significant differences in increase in matched paired objective test (30 questions),

increase in matched paired students' attitude, overall pre-test and post-test scores, and overall pre-test and post-test attitudinal scores. Previous research in education may have suggested that differences in socialization may affect male and female students' learning. However, the results in this study indicated that the JA BizTown curriculum can improve the students' knowledge in Finance and Economics for both male (average increase of 4.5) and female (average increase of 4.4) students with no significant difference in score increase between gender groups.

Analysis of ethnicity

Hypothesis 2 suggested that there would be no difference in students' understanding of and attitude towards economics and financial concepts between Black and White groups. With respect to ethnicity, students were asked to indicate their race by answering a Yes/No for being in a certain racial group or partly in that group. There are 41% of students who indicated White or part White, 50% Black or part Black, 10% Hispanic or part Hispanic, 6% Asian or part Asian, 11% American Indian/Alaskan Native, and 12% Others. Three comparisons were made focusing on those in the White, Black, and Hispanic groups in terms of increase in matched paired objective test (30 questions) and increase in matched paired students' attitude. The MANOVA result indicated no statistically significant improvement in objective test scores but significant reduction of attitudinal scores between those who are White or part White and those who are not White at all. When comparing students who are Black/part Black vs. not Black, there is statistically significant higher increase in objective test scores in the not Black group (5.24) than the Black/part Black group (4.1) but significant reduction in attitudinal scores. The same analysis was applied to the Hispanic group but the results were not significant. The findings give support to the hypothesis that different racial groups have different students' understanding of economics and financial concepts. The reduction in attitudinal scores, however, is not expected.

Analysis of Previous JA Participation

Hypothesis 3 suggested that previous JA participation is related to students' understanding of economics and financial concepts. Students with previous JA experience vs. those with no JA experience were compared in terms of increase in matched paired objective test (30 questions), increase in matched paired students' attitude, overall pre-test and post-test scores, and overall pre-test and post-test attitudinal scores using MANOVA. . The results indicated insufficient evidence to find statistically significant differences in increase in matched paired objective test (30 questions) but students' attitude scores were reduced more in the group with previous JA participation than the one without previous JA participation. 359 students have indicated that they have participated in other JA programs before JA BizTown. The number is much higher than expected which could mean the students may not understand the question completely.

Analysis of factors affecting students' understanding of economics and financial concepts

Hypotheses 4 to 7 suggested students' understanding of economics and financial concepts could be related students' attitude towards education, the number of teachers involved in teaching the curriculum, time spent on lessons, and teachers' effectiveness. Correlations among these variables are examined. With regard to increase in objective test scores, the only factor having significant relationship is students' perception of teachers' effectiveness of teaching the JA BizTown lessons. The more positive is a student's perception of teacher's effectiveness, the higher the increase of test score is expected. The data did not support relationship between increase in test scores and students' attitude towards education, the number of teachers involved in JA BizTown lessons, and the time spent on lessons. When analyzing the changes in students' attitude before and after JA BizTown curriculum, the results indicated significant relationship with students' attitude towards education, the number of teachers involved in teaching the curriculum, time spent on lessons, and teachers' effectiveness. As expected, when the students perceived more positively in the amount time spent on lessons and teachers' teaching effectiveness, the changes in students' attitude before and after JA BizTown curriculum were more positive. Notable in the pattern of positive relationships we found is a pattern of unexpected *negative* relationship between increase in attitude scores and students' attitude towards education and number of teachers involved in JA BizTown lessons. Certainly, the literature leads to the expectation that students attitude towards education and the number of teachers involved would be found with more "positively" related to attitude improvement towards Economics and Finance. Yet these results suggested the contrary. Note that changing students' attitude is a long term process. Perhaps in the future study, we can track students' general attitude towards learning and their confidence levels over several years using a program that lasts not just a semester in order to bring deep rooted attitudinal changes.

		Attitude towards education	Number of teachers teaching JA lessons	The amount of time your class spent on each lesson is just right	Teacher(s) has done a good job teaching the BizTown lessons
Increase in objective test scores	Pearson Correlation	.024	-.022	.017	.062(*)
	Sig. (2-tailed)	.442	.496	.589	.049
	N	1024	1002	1005	1000
Increase in student attitude scores	Pearson Correlation	-.360(**)	-.224(**)	.098(**)	.356(**)
	Sig. (2-tailed)	.000	.000	.003	.000
	N	951	943	947	942

LIMITATIONS

In order to measure the students' knowledge of economic and financial concepts, the students need to complete the pretest and post-test along with the Student Pre- and Post-Program Surveys in an efficient and effective manner. There are a few concerns. The first is the fact that some schools had missing data due to student absences and lack of participation from teachers. The second concern focuses on the massive amount of data entry which could lead to data entry errors and delay in data entry. To address the first concern, JA staff will have to work diligently with school administrators and teachers by providing them pre-program training, constant communication via school visits, emails, and phone calls. This will reduce the amount of missing data. To address the data entry problem, JA will have to have additional staff to perform data processing and entry duties. This will ensure data integrity and validity.

REFERENCES

"A nation still at risk." (2008) *American School Board Journal* 195(6), 7.

Americans for Consumer Education and Competition. *Nation's high school seniors fail at finance fundamentals.* Retrieved April 17, 2002, from http://www.acecusa.org/releases/010315.asp

American Savings Education Council. *1999 Youth & Money: Results of the 1999 youth and money survey.* Retrieved July 3, 2003, from http://www.asec.org/media/pr103.htm.

Auger, R. W., A.E. Blackhurst, & K. H. Wahl. (2005). The Development of Elementary-Aged Children's Career Aspirations and Expectations. *Professional School Counseling*, 8(4), 310-321.

Bandura, A., C. Barbaranelli, G.V. Caprara, & C. Pastorelli (2001). Self-Efficacy Beliefs as Shapers of Children's Aspirations and Career Trajectories. *Child Development*, 72(1), 187-206.

Becker, W.E., & Walstad, W.B.(1987). *Econometric modeling in economic education research*. Boston, MA: Kluwer-Nijhoff.

Black, S. (2009) An Investment in Literacy. *American School Board Journal,* 196(2), 44-45.

Finkel, Ed. (2010) A Bull Market for Financial Literacy: The Great Recession compels school districts to teach dollars and sense. *District Administration* 46(8), 68-74.

Varcoe, K., A. Martin, Z. Divitto, & C. Go. (2005). Using A Financial Education Curriculum For Teens. *Journal of Financial Counseling & Planning* 16(1), 63-71.

APPENDIX 1: STAGES OF JA BIZTOWN IMPLEMENTATION

Planning: JA Capstone Program staff will meet with each education provider to secure a signed, written Memorandum of Understanding (MOU) outlining the roles and responsibilities of JA and educational providers. JA agrees to provide curriculum materials, in-class assistance and consultation, simulation site and materials, educator/volunteer training, and staff assistance during simulation. The teachers agree to attend training, accept and use curriculum materials as directed in the training, inform adult volunteers and parents of the date and time of the simulation, and enforce behavioral expectations of the students. The MOU also states school specific accommodations as necessary. The agreement is signed by the school principal, lead

teacher and Capstone Education Manager. During this phase, the lead teacher will also work with the Capstone Education Manager to complete a written schedule that complies with the program requirements and will allow JA to predict when participating students are expected to attain certain lesson goals. Planning will be completed up to two months prior to beginning the in-class program.

Training: Teachers and volunteers will be required to attend a training session to participate in the program. Volunteers are specifically trained to facilitate the simulation exercise and to not complete tasks for the students. The 2-hour session for teachers is held no later than one week prior to the beginning of the in-class units, and the 1.5-hour session for volunteers is held one week prior to the on-site simulation and scheduled in the evening to accommodate working parents. The schools are asked to recruit *up to 14 parents and others* to volunteer for the simulation. In the spring 2009 test phase, schools were successful in recruiting 5-6 parent volunteers per simulation. The trainings are compressed into intense sessions in sensitivity to the highly stressed school system, its educator demands, and the demands of working parent volunteers.

Classroom Instruction: The Capstone Education Manager will introduce the first lesson and embed the program's relevance in a real-life context, enhance student excitement, underscore learning expectations necessary for participation in the simulation, and administer the pre-test. Teachers will teach all other lessons independently, and the Capstone Manager will rejoin the teacher for debriefings. Lessons include active learning sessions in economics and the free enterprise system and incorporate state mandated standards in English, language, mathematics and social studies. Taught in a minimum of 18, 45-minute in-class sessions (divided between five distinct units), the lessons are designed to minimize preparation time. Optional learning activities also provide opportunities for students to strengthen skills through practice in language arts, math and social studies. Teaching strategies include cooperative learning, concept definitions, role play, writing, webbing, creative thinking, and others to engage all students in the learning process. Each class will be required to meet benchmarks that show an overall GPA of 2.5 or higher on each unit test and other graded assignments. The Capstone Manager will contact the teachers through all available channels of information to ascertain that benchmarks reflecting required minimum GPA at each level were reached **at the end of each unit**, and the manager will use all communication means available to obtain data that proves in-class progress, including copies of recorded grades. The manager will remind teachers of learning expectations before the anticipated conclusion of each unit, based on the pre-set schedule developed in the Planning Stage, so that if classes that are lagging behind expectations, it is still possible for teachers to take corrective measures. Classes that do not reach the required benchmarks in all units will not be allowed to participate in the simulation.

Simulation: Classes demonstrating positive outcomes will participate in a full-day, simulation exercise in a highly technology based learning center, that is *only* available to our schools through JA BizTown, where students use their learned classroom skills to master tasks

involved in citizenship, business management and personal finance. Two weeks before the scheduled site simulation, JA staff will re-confirm site simulation dates, any appropriate accommodations, and times with all participating school principals and educators. JA staff will make a second re-confirmation two school days before the scheduled simulation. In preparation for the simulation, students apply for jobs, interview for positions, and review their assigned job functions. In the simulation, students role play as employees in one of 14 businesses (facilitated by volunteers and educators). Students are divided into business teams charged with operating a quality business and practice their learned classroom skills, including time and money management, banking, entrepreneurship, consumerism, and making wise choices using money. In the simulation exercise, students earn two mock paychecks for their labor, and they manage personal checking and savings accounts, deposit earnings and withdraw funds from the bank, supervise themselves and others, and monitor business/personal expenses. JA will supply and prepare the simulation site for each simulation experience. Each business CEO and CFO must supervise their employees, pay business expenses and pay down a business loan. For example, Entergy New Orleans installed real meters that the students, employed as Entergy workers, must read, determine electricity usage, and bill other businesses for the expense. Students, on their lunch break, may "purchase" their lunch in the restaurant assisted by the restaurant waiters and manager – all 10-12 year olds. On morning and afternoon breaks, they may "shop," make spending choices based on income availability and check their stock portfolios for rises and falls in the market. Students employed in the TV studio produce business ads, which are broadcast on closed circuit monitors, and produce a DVD of the day that goes back to the classroom with the teachers. The newspaper office employees complete a publication, including stories from interviews and student produced digital photos. JA equips, supplies, and prepares the center for each experience, which accommodates 50 to 100 children for each simulation.

In-Class Follow Up: The program concludes with student debriefing lessons led by a JA staff/educator team to ensure a well-rounded learning experience, allowing students to reflect on their experiences and confirm the link between classroom learning and their future plans and goals. Students will collectively evaluate student team performances, describe their personal experiences in a business letter and identify what they did well and what they would change if their businesses were to continue. Students are asked to demonstrate their knowledge by explaining the circular flow of economic activity, describing how citizens use financial institutions, and describing how citizens work within a quality business. The post-test is administered at this time. Each child will write an essay on his or her thoughts on a future career choice, research the education required to achieve this choice, and reflect on his or her individual role as part of a community and the free enterprise system.

THE BENEFITS OF ASYNCHRONOUS DISCUSSION IN A HYBRID COURSE: EVIDENCE FROM A LARGE ENROLLMENT ECONOMICS COURSE

Travis Roach, Texas Tech University

ABSTRACT

As improvements in technology continue to be integrated within the collegiate classroom it is important to study the benefits, or costs, that are associated with adopting new pedagogical practices. This paper focuses on the role that asynchronous discussion can play in furthering student development within a hybrid economics course. Specifically, this paper finds that encouraging online discussion of articles, podcasts, and videos that are related to course material results in better academic performance.

Keywords: *current events, blended-learning, asynchronous-learning, asynchronous-discussion, hybrid course*

INTRODUCTION AND BACKGROUND

The prestigious ivory towers that come to mind when one thinks of taking and attending college classes are slowly being replaced by their digital counterparts. While there is still a place for the time-honored tradition of lecturing and conventional face-to-face teaching methods, pedagogical research has begun to highlight the very interesting world of technology in the classroom. Blended learning, as it is often referred to, is the conscientious integration of online learning experiences with established face-to-face practices. Garrison and Kanuka (2004) conclude their article on the transformative potential of blended learning by saying that, "blended learning can begin the necessary process of redefining higher education institutions as being learning centered and facilitating a higher learning experience" (Garrison & Kanuka, 2004). Blended learning has a special connection to economics education because of the ease at which economic principles can be applied to news from around the world. As many economists have taken to the internet to write blogs for both their classrooms and the public, more classes have started to tend towards the "hybrid" course format; perhaps without even meaning to do so. Despite the volume of literature focusing on the pedagogical potential of online learning in a blended or "hybrid class", little has been done to test the efficacy of these practices. This paper contributes to the literature by testing whether or not online asynchronous discussion of current events truly furthers student learning of core concepts. This is achieved by considering data on 311 students enrolled in principles of microeconomics across two semesters and testing whether

or not contributing to an online discussion improved performance on a standard subject test. The data show that asynchronous learning has in fact improved student cognition when test grades are considered as the dependent variable, while controlling for other factors that may affect student performance.

The term "hybrid course" has been used by many authors to distinguish courses that use both face-to-face and distributed (distance) learning tactics. For the purpose of this paper a hybrid course is one which combines face-to-face instruction with computer-mediated instruction (Graham, 2006; Reay, 2001; Rooney, 2003; Sands, 2002; Young, 2002). The practice of using technology in the classroom has been the study of many authors (Lin, 2007; Martyn, 2003; Massoud, et al., 2011). In fact, many have found that in a technology-rich learning environment, learner-centered and active-learning techniques are more commonly used (Graham, 2006; Hartman, Dziuban, and Moskal, 1999). The increased use of active-learning due to technology is a boon to cognitive development *a la* Bloom (1956) because blended learning environments are able to foster interaction and allow students to connect with the learning materials and each other. In Bloom's seminal work, which has since sparked entire areas of educational research, the cognitive domain is separated into six tiers – knowledge, comprehension, application, analysis, synthesis, and evaluation. Active-learning activities and exercises are easily able to target the upper tiers of Bloom's taxonomy because they compel students to become active participants and *apply* knowledge they have learned and *evaluate* outcomes. Yamarik (2007) studied the use of active learning in the economics classroom and found that students who were exposed to an active-learning environment performed better on tests than those who did not. Active-learning does not need to be confined to classroom instruction; it can easily be utilized in an online environment. For instance, many authors have found that when asynchronous text-based discussion is used, students can carefully reflect on and provide evidence for their claims. The resulting discussion contribution allows for deeper, more thoughtful reflections on the part of the learner (Graham, 2006; Mikulecky, 1998; Benbunan-Fich & Hiltz, 1999).

Recently, Bloom's taxonomy and it's applications to active and cooperative learning have been discussed in the context of technology-rich environments (Kenney & Newcombe, 2011; Chang & Fisher, 2003). In fact, Kausar et al. (2003) developed a study that rated computer assisted learning against lecture based learning in terms of Bloom's taxonomy. The authors found that computer assisted learning was indeed superior to traditional lecture based teaching.
A key to the success of a hybrid course is the integration of online materials within the lecture. Even in the early years of the internet teachers began realizing the many different learning opportunities that were available (Simkins, 1999). In-class discussion is an excellent way to incorporate online materials. Allowing students to apply the stories they read or hear to classroom concepts during a discussion certainly approaches the top tiers of Bloom's taxonomy - synthesis and evaluation.

In addition to classroom discussion, which is regarded by Brookfield (2005) as a tool for fostering critical thinking skills, online discussion is an excellent way to develop critical thinking skills and include different learning styles and personalities. Online discussions which progress

at the pace that students post to a forum or discussion board is often referred to as "asynchronous learning." Many authors have noted that knowledge sharing in an online discussion has the potential to improve student learning (Brewer & Brewer, 2010; Kienle, 2009). Bassett (2011) studies how students view asynchronous learning and finds that "the collaborative nature of the online discussions facilitated an inclusive learning experience for all students." Also speaking to the inclusive nature of online learning, Miyazoe and Anderson (2011) found that "online writing assignments using pseudonyms can be an effective teaching strategy that induces higher online participation, especially among students who are hesitant to participate in a traditional classroom setting." Similarly, Gerbic (2010) finds that timid students who may not want to discuss in a classroom setting are less inhibited when discussing online. Gerbic also points out that international students view online discussion as a safer environment in which to participate. The potential for attracting students who would normally not contribute in a classroom situation helps ensure that multiple learning styles and personality traits are reached

Speaking to economics in particular, a few authors have taken the task of assessing online learning and hybrid classes. Harter and Harter (2004) consider a similar topic as this paper and find that teaching with technology does not improve student scores on tests. Gratton-Lavoie and Stanley (2009) find that strictly online courses, as compared to traditional or hybrid courses, may have insignificant and sometimes negative impacts on student learning. On the other hand, Navarro and Shoemaker (2001, 2000b) found that online learning benefitted students.

COURSE AND ONLINE COMPONENT DESCRIPTION

This study was conducted using data from students in an introductory-level principles of microeconomics course. The course design is such that students receive face-to-face instruction for three hours a week, but also use online resources outside of class in the form of "current events" that can be found on a blog prepared by the author.

The Current Events Blog

The current events blog of online resources was used to compliment course reading by linking to various news articles, podcasts, and videos that applied to the material currently being discussed. For brevity's sake, the various different media that students used will be referred to as "stories" here forward. In fact, designating articles, podcasts, and videos as stories truly describes the aim in this project of using outside resources to compliment the teaching of a concept. Creating a narrative to accompany complex theories or concepts is a form of experiential learning (Dalton 2011). Itin (1999) defines experiential learning as any instance in which an individual derives meaning from personal direct experience. The importance of experience is being considered in other scenarios than education. Pine and Gilmore (1999) describe the "experience economy" as being the next step in the evolution of consumer

preference. If education is to be tailored to the needs of today's consumers, then experience is in high demand. In a related article, Gilmour (2003) pointed out that the average consumer is willing to spend more money on compelling experiences that connect and inspire them in a personal and unforgettable way than a bland alternative. In order to incorporate the idea of the experience economy within Itin's definition of experiential learning, the hybrid-course variant proposed is that experiential learning occurs when an individual derives meaning from indirect experience through online media. Moreover, experiential learning is not only what we as educators ought to sell to our students; it is what they want to buy from us.

The online stories that are used are not to be viewed as homework in the traditional sense because they are introduced to students in such a way that online instruction occurs. Each story can be found on a blog that is prepared by the instructor. Along with the link to the story, each post is accompanied with a short discussion of the story and a few open-ended questions that are intended to help start online discussions or direct the student's attention. The following is an example of a current events post from the second semester that received numerous posts to the discussion thread:

> *"Here's a quick supply and demand problem. The officials at Foxconn (The main hub of Apple manufacturing) are going to increase the wages paid to their workers. What do you expect to happen to the price of iPhones and iPads? I assume that Foxconn's decision to increase wages came from pressure by Apple who has recently been under scrutiny for how foreign labor has been treated. Do you think, though, that Americans would still be pressuring Apple to pay the people who manufacture their products more if they knew what this would do to the price of Apple products?"*

In response one student posted,

> *"I think that, once the issue of inhumane work conditions are raised, people tend to hold that banner over a large corporation without the realization of what that cost will be. However, I don't think the insistence for better care of workers will stop when the prices go up. I'm not sure most people will even connect the two. Instead, they will become two tallies against Apple, instead of a cause and effect as Apple tries to fix the former by bringing about the latter."*

This post in turn solicited many other thoughtful comments. It is clear from the above comment that when students are given the time and relaxed environment of online learning they can consider a basic concept like supply and demand at very high levels. In terms of Bloom's taxonomy, this student has taken a basic supply and demand problem and answered it at the evaluation level.

Another aspect of the course design is that materials are easily available to students by mobile device or tablet computer. By designing the website and blog in such a way that it is

easily viewed on a small mobile device, students are able to read, listen, or watch the assigned current events at their leisure.

RESEARCH DESIGN & EMPIRICAL ANALYSIS

In order to separate the effects that asynchronous interaction has on student performance several attributes of the course had to be kept similar between the two semesters. For both semesters the class met on Mondays, Wednesdays, and Fridays at close to the same time. Also, both classes had a large class size with 175 students and 136 students for the fall and spring, respectively. The amount of posts and the type of content posted to the current events blog was similar for both semesters as well. In order to enforce that students follow the blog, both semesters of students were given a quiz over the current events material. Although the stories associated with the quiz were different for each semester, the format of the quiz was particularly alike between semesters. Finally, to accurately control for any bias caused by students leaving the class at mid-semester only the time span from the beginning of the semester to the first test has been considered.

The difference between both semesters was that students in the spring semester were told at the beginning that they would need to contribute to the blog by posting one comment and one response to a comment for the story of their choice. To motivate students to fulfill this requirement a total of 10 points were allocated on each quiz covering the current events. In other words, if students failed to make comments the best grade that could be received on the current events quiz was a 90. The comments were graded based solely on participation, though students were told that each comment should be at least 2-3 sentences in length. Finally, students in the spring semester were asked to post under the pseudonym of their student ID number. This allowed for all posts to be anonymous to others. Interestingly, some people chose not to remain anonymous and posted using their names. Of the students who chose to identify themselves they were mostly male, and if it was done by a female student she would use her last name instead of her full name. This finding motivates the use of gender as an explanatory variable in the following section, although we will see that there are not any differences due to gender in the returns to discussing asynchronously.

Because everything, except for the emphasis that students post to the current events blog in the spring semester, was held constant we can deduce that improvements in student performance are a result of asynchronous discussion participation. As a measurement of student performance, test scores on the first test are used. For both semesters this test was standardized in such a way that they were the same length with very similar, and in many cases identical, multiple choice questions. Leaking of the test was prevented as well.

Empirical Analysis

A priori, it is expected that more interaction and discussion of the course and its concepts via asynchronous discussion will translate to higher grades on tests. To test this hypothesis the following linear regression model is considered,

$$Grade_t = \beta_0 + \beta_1 Business_t + \beta_2 Attend_t + \beta_3 Under_t + \beta_4 AsyncLearn_t + \varepsilon_t$$

where: Grade is the student's test grade on the standardized first test; Business is a dummy variable measuring if the student is a business major or not, Attend measures the students attendance percentage; Under is a dummy variable measuring if the student is a freshman/sophomore; AsyncLearn is the amount of comments to articles that the student made. Summary statistics for all variables can be seen in table one, and correlation coefficients can be seen in table two.

Table 1 – Summary Statistics

Variable	Mean	Median	Std. Dev.
Test1	71.9191	72.2222	14.2666
AsyncLearn	0.797428	0.000000	1.11332
Business	0.450161	0.000000	0.498312
Attend	78.6446	88.8889	24.1275
Under	0.755627	1.00000	0.430407
Notes: 311 observations			

Table 2 – Correlation Coefficients

Test1	AsyncLearn	Business	Attend	Under	
1.0000	0.1762	0.1397	0.2925	-0.0165	Test1
	1.0000	0.0291	0.2648	0.0795	AsyncLearn
		1.0000	0.1403	0.1857	Business
			1.0000	0.0724	Attend
				1.0000	Under

The course is one of a few options in the core curriculum for arts and sciences majors and is mandatory for all business majors. Hence, the concentration of business majors in the class is high. The variable, Business, was introduced not only to capture the large amount of business majors, but also to proxy for innate ability. Ability is clearly not observable, but by separating the autonomous effect of being a business major it is assumed that those who are inclined to "think economically" have been accounted for. Along the same lines, the variable for attendance was included to proxy for student motivation. Gratton-Lavoi and Stanley (2009) study the effects of online learning for microeconomics students as well. A concern that their paper raises is that selection bias may occur when choosing the method of instruction – online, hybrid, or traditional upon registration. For this study, however, students were not aware of any difference in

instruction method when registering for the course. Hence, selection bias, of this nature, is not perceived to be a problem.

A few other variables were considered for the regression analysis including a dummy variable for gender, and a dummy variable for whether or not the student had taken any economics course at the collegiate level previously. The variable for gender was omitted in the final analysis because it did not aid the predictability of the model and was statistically insignificant. This too departs from Gratton-Lavoi and Stanley (2009). In their study they find that male students do on average 7.5 points better than their female classmates in a hybrid class. The variable that measured past experience in collegiate economics courses was also omitted because it was insignificant and not necessarily needed theoretically. 72 students had been previously enrolled in an economics course out of the 311 observations, and most of them were repeating the course in order to replace a low grade. It cannot be determined, then, how much was gained from the student's previous experience in a collegiate economics course.

It is supposed that asynchronous discussion is beneficial to the student but not confirmed through previous studies. If the coefficient for AsyncLearn is positive, and significant, this will signify that asynchronous discussion has in fact increased the student's comprehension of subject material. If it is negative, then asynchronous discussion has been detrimental to subject material comprehension. Because business is a closely related field to economics, it is expected that business majors will perform better on tests than their non-business classmates. The coefficient for Attend is expected to be positive implying that the more often a student attends class, the better their grade will be. Finally, it is expected that underclassmen will not be as successful on tests as older students are.

To estimate the linear model proposed above, the method of ordinary least squares (OLS) is used. Many authors consider the potential for selection bias with OLS results. In this study, selection bias could occur because students who would benefit from discussing material online may not be accounted for because all second semester students were required to make two comments. Some students, though, contributed more than the mandatory amount of comments in the spring, or made voluntary comments in the fall semester. For these students, the benefit to grades could be biased because they have a pre-disposition to this type of learning. In consideration, a Heckit model was estimated using a binary variable that measured whether or not a student contributed more than the mandatory amount. Heckman (1979) originally proposed this methodology to account for sample selection bias when the wages of workers are considered. This bias is overcome, if it exists, by estimating two equations: a selection equation - did the student contribute more than the mandatory amount; and an outcome equation - test grade. The variable, spring, is used in the selection equation because the amount of effort required to post an extra comment is very low for spring semester students. In other words, while a student is already on the blog studying and posting comments it is not odd for the student to post again in the midst of many other comments whereas a post by a student in the fall semester would likely be the only comment associated with that story. In the results for the Heckit model

(found in table 2), lambda, which can be interpreted as the correlation between the error terms in the grade equation and the selection equation, was positive but insignificant. Hence, selection bias is not of concern for this paper and the OLS regression results can be interpreted in the traditional way. In order to estimate the Heckit model the traditional two-step process was used. Mostly, this is due to the strong assumptions necessary for maximum likelihood estimation of the Heckit model (Greene, 2008; Wooldridge, 2002). A detailed table of the Heckit model results can be found in table three. Regression results from the OLS model can be found in table four.

Table 3 – Heckit Estimation

Outcome equation					
Variable	**Coefficient**	**Std. Error**	**z**	**p-value**	
const	52.4467	13.9774	3.7522	0.00018	***
Business	-9.12105	6.68098	-1.3652	0.17218	
Attend	0.241546	0.15965	1.5130	0.13029	
Under	-4.35489	5.28507	-0.8240	0.40994	
AsyncLearn	2.51429	2.57365	0.9769	0.32860	
lambda	9.37771	9.21434	1.0177	0.30881	
Selection equation					
const	-1.78329	1.03191	-1.7281	0.08396	*
Business	-0.830667	0.742058	-1.1194	0.26297	
Attend	-0.00600933	0.0137039	-0.4385	0.66102	
Under	-0.6031	0.576632	-1.0459	0.29561	
AsyncLearn	3.11023	0.524648	5.9282	<0.00001	***
Spring	-5.76886	1.24267	-4.6423	<0.00001	***
sigma	10.78168		rho	0.869782	

Notes: Heckit - 2-Step method, QML standard errors, 311 Observations, 289 censored observations, *$p < .05$; **$p < .01$; ***$p < .001$

Table 4 – Regression Results (OLS)

Variable	**Coefficient**	**Std. Error**	**t-ratio**	**p-value**	
const	59.027	3.16827	18.6307	<0.00001	***
Business	3.23862	1.6569	1.9546	0.05154	*
Attend	0.150909	0.0363452	4.1521	0.00004	***
Under	-2.13974	1.78182	-1.2009	0.23073	
AsyncLearn	1.42597	0.688895	2.0699	0.03930	**
R-squared	0.110136	Adj. R-squared	0.098466		
F(4, 305)	9.928666	P-value(F)	1.45e-07		

Notes: Heteroskedasticity corrected standard errors. *$p < .05$; **$p < .01$; ***$p < .001$

The estimated values for all variables are consistent with the expected sign for each variable. According to the estimates, business majors do in fact perform better on tests than non-business majors by approximately 3 points. Also as expected, the attendance rate is very significant in determining a student's grade. The estimates show that if a student increases their

attendance by 10%, their test grade will increase by about 1.5 points, all else constant. The main variable of interest is AsyncLearn which is both positive and significant at the 5% level. The estimates show that for each additional comment on the current events blog a student raises their grade by about 1.4 points. Or in other words, if the only difference between two students is that one student was more involved in discussing course related material outside of class, i.e. posting comments to the blog, the student who discussed asynchronously will perform better on the test.

The R-squared statistic for the OLS model is approximately .11 which is on the low side of acceptable. For this reason, a least absolute deviation (LAD) model is estimated and presented in table three. Least absolute deviation regression is an estimation technique that is more robust than OLS when data that have many observations at the low or high end are considered. As one might expect in a large freshman level class, test grades are approximately normally distributed but with a "non-normal" amount of observations on the low end. The data in this study are no exception.

Table 5 – Regression Results (LAD)

Variable	Coefficient	Std. Error	t-ratio	p-value	
const	59.9722	4.37352	13.7126	<0.00001	***
Business	3.0000	1.82204	1.6465	0.10069	
Attend	0.131944	0.0463575	2.8462	0.00472	***
Under	-1.86111	2.04112	-0.9118	0.36259	
AsyncLearn	2.54839	1.01307	2.5155	0.01240	**
Sum absolute resid.	3175.863	Sum squared resid.	56496.33		
Notes: *$p < .05$; **$p < .01$; ***$p < .001$					

The estimates from the LAD model are quite similar to the OLS results. Still, as attendance increases, so too does the expected test grade. Business majors are still predicted to have higher test scores on average, but this can only be said with about 90% confidence. The measure for asynchronous learning, however, has increased both quantitatively and statistically. Before, each additional discussion post garnered an expected 1.4 points on the test, but the LAD model predicts that each additional post will increase a student's test grade by about 2.5 points. To put this in context, if a student completed 2 discussion posts their test grade is expected to be about 5 points higher than students who did not post. Hence, both regressions support the hypothesis that asynchronous discussion of course-related online materials increase student performance on standard subject tests.

CONCLUSION

This paper has shown that asynchronous discussion of course-related materials has in fact improved student comprehension of course material. This is likely due to the higher levels of

thinking that can occur when students interact with each other and go beyond answering simple multiple choice type questions. By actually applying and evaluating concepts learned in class to real life examples, and by furthering this knowledge with original contributions to a comment thread, students have elevated their learning of lecture material. The results of this paper are encouraging to the development of hybrid classes, but should be met with a little reservation. Because asynchronous learning has worked in this principles of microeconomics classroom does not mean that it will necessarily work with other subjects or fields of economics. The method of online dissemination, the types of materials used online, and the environment in which students discuss are all major variables to consider. Hence, more study on the efficacy of asynchronous discussion in a hybrid class is still needed to fully support its use pedagogically.

REFERENCES

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & D.R. Krathwohl. (1956). *Taxonomy of educational objectives: the classification of educational goals; Handbook I: Cognitive Domain* New York, Longmans, Green.

Brewer, P. D. & K. L. Brewer. (2010). Knowledge management, Human Resource Management, and Higher Education: A theoretical model. *Journal of Education for Business*, *85*, 330–335.

Brookfield, S. 2005. *Discussion as a Way of Teaching of Teaching: Tools and Techniques for Democratic Classrooms*, San Francisco, CA: Jossey-Bass.

Chang, V. & D. Fisher. (2003). The Validation and Application of a New Learning Environment Instrument for Online Learning in Higher Education. In M. Khine, & D. Fisher (Eds.), New Jersey *Technology-Rich Learning Environment.*World Scientific Publishing.

Gerbic, P. (2010). Getting the blend right in new learning environments: A complementary approach to online discussions. *Education and Information Technologies, 15,* 125-137.

Gilmour, J. H. 2003. Frontiers of the Experience Economy. In *Batten Briefings: From the Darden Graduate School of Business Administration*. Charlottesville, VA. (http://www.strategichorizons.com/documents/BattenBriefings-03Fall-FrontiersOfEE.pdf Accessed online 4/7/2012)

Graham, C. R. 2006. Blended Learning Systems: Definition, Current Trends and Future Directions. In C. J. Bonk, & C. R. Graham (Eds.), *The Handbook of Blended Learning: Global Perspectives, Local Designs*. San Francisco, CA

Garrison, D. R., & H. Kanuka. (2004). Blended learning: Uncovering its transformative potential in higher education. *Internet and Higher Education*, 7, 95-105.

Gratton-Lavoie & C. D. Stanley. (2009). Teaching and Learning Principles of Microeconomics Online: An Empirical Assessment. *Journal of Economic Education*, *40*(1), 3-25.

Greene, W. 2008. *Econometric Analysis* 6th Ed. Prentice Hall, New Jersey.

Harter, C. L. & J. F. Harter. (2004). Teaching with technology: Does access to computer technology increase student achievement? *Eastern Economic Journal,* 30 (Fall), 507–14.

Heckman, J. (1979). Sample Selection Bias as a Specification Error. *Econometrica,* 47 (1), 153–61.

Itin, C.M. (1999). Reasserting the philosophy of experiential education as a vehicle for change in the 21st century. *Journal of Experiential Education,* 22, 91–98.

Kausar, T., Choudhry, B. N. & A. A. Gujjar. (2008). A Comparitive Study toEvaluate the Effectiveness of Computer Assisted Instruction Versus Classroom Lecture for Computer Science at the ICL Level. *The Turkish Online Journal of Educational Technology,* 7(4).

Kenney, J. & E. Newcombe. (2011). Adopting a Blended Learning Approach: Challenges Encountered and Lessons Learned in an Action Research Study. *Journal of Asynchronous Learning* 15(1).

Kienle, A. (2009). Intertwining Synchronous and Asynchronous Communication to Support Collaborative Learning—System Design and Evaluation. *Education and Information Technologies*, *14*(1), 55–79.

Lin, H. (2007). Blending Online Components into Traditional Instruction: A Case of Using Technologies to Support Good Practices in Pre-Service Teacher Education. *Journal of Instructional Delivery Systems,* 21(1), 7-16.

Martyn, M. (2003). The Hybrid Online Model: Good Practice. *EDUCAUSE Quarterly,* 1, 18-23.

Massoud, A. et al. (2011). Using Blended Learning to Foster Education in a Contemporary Classroom. *Transformative Dialogues: Teaching & Learning Journal* 5(2).

Navarro, P., & J. Shoemaker. (2000a). Performance and Perceptions of Distance Learners in Cyberspace. *American Journal of Distance Education,* 14 (October), 15–35.

Navarro, P., & J. Shoemaker. (2000b). Policy Issues in the Teaching of Economics in Cyberspace: Research Design, Course Design, and Research Results. *Contemporary Economic Policy,* 18 (July), 359–66.

Pine, B. J. & J. H. Gilmore. (1999). *The Experience Economy: Work is Theater & Every Business a Stage: Goods & Services are No Longer Enough*, Boston, MA: Harvard Business Review.

Simkins, S. P. (1999). Promoting Active-Student Learning Using the World Wide Web in Economics Courses. *Journal of Economic Education,* 30 (3), 278–91.

Wooldridge, J. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge: MIT Press.

Yamarik, S. (2007). Does Cooperative Learning Improve Student Learning Outcomes? *Journal of Economic Education,* 38 (3), 259-77.

A STUDY OF STUDENTS' VIEWS OF MARKET FAIRNESS

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

ABSTRACT

Although this study was prompted by the recent "Occupy" movements, the paper utilizes two studies on the role of "fairness" in economic situations: one by Kahneman, Knetsch, and Thaler (1986b) and a second by Shiller, Boycko, and Korobov (1991). This study employs eight (8) scenarios used in either the Kahneman et al. or Shiller et al. studies to investigate the existence of differences in the perception of the fairness of markets along both gender lines and major field of study. Data were gathered in an anonymous in-class survey of first-year university students. Overall, male students generally had a more favorable impression of markets than females. Surprisingly, the results of the Business and Non-Business students were mixed on the fairness of pricing.

INTRODUCTION

The typical introductory Economics text discusses "The Three Questions" that any society must address: 1) *What* goods are to be produced?; 2) *How* are those goods to be produced?; and 3) *For Whom* are the goods produced? When it comes to discussing the third question, the typical instructor in the United States focuses on the role played by markets. However, Colander (2003) contends that the current majority of principles textbooks "excludes discussion of a broader set of failures-of-market outcomes: failures in which the market is doing everything it is supposed to be doing, but society is still unhappy with the result" (p. 83). In today's society, recently highlighted by the various "Occupy" movements, many people view the issue as whether the market is "fair", or at least perceived to be "fair".

Kahneman, Knetsch, and Thaler (1986b) studied the role played by the perception of fairness in explaining economic situations. Specifically, the two primary objectives of the study were to identify community standards of price fairness and the possible implications of the rules of fairness for market outcomes. The authors created 18 scenarios and collected data over 14 months in a series of telephone interviews of randomly selected residents of Toronto and Vancouver. The respondents were composed of an approximately equal number of both males and females, were read no more than five of the 18 scenarios, and were asked to respond to each

scenario with the categories "Completely Fair", "Acceptable", "Unfair", and "Very Unfair". In the article, the two favorable responses and the two unfavorable responses were collapsed into the categories of "Acceptable" and "Unfair" to indicate the proportions of respondents who judged the action acceptable or not. Kahneman et al. found respondents had a strong aversion to price rationing (resulting in some price friction), consumers were more tolerant of price changes resulting from a changing cost structure (than price changes attributed to demand considerations), and a general dislike for the use and exploitation of market power. The authors concluded:

> *The findings of this study suggest that many actions that are both profitable in the short run and not obviously dishonest are likely to be perceived as unfair exploitations of market power. Such perceptions can have significant consequences if they find expression in legislation or regulation (Kahneman, Knetsch, and Thaler, 1986b, pp. 738-739).*

Gorman and Kehr (1992) used 16 of the 18 scenarios developed by Kahneman et al., and created six additional contrasting scenarios. The authors used a total of 22 scenarios in a survey mailed to randomly selected business executives. The authors' intent was to determine whether a sample of business executives would respond to the scenarios in a different manner than the general population sample by Kahneman et al. With 154 business executives responding, the authors concluded that business executives have a different perception of market fairness than the general public. Specifically, the business executives responding to the survey were less inclined to judge the profit-maximizing behavior as unfair.

Shiller, Boycko, and Korobov (1991) designed 36 scenarios pertaining to "*fundamental* parameters of human behavior related to the success of free markets" (p. 386, italics in original). The 36 scenarios were partitioned into three sets of 12 and administered in a series of telephone interviews to residents of Moscow and New York City. The responses were categorical in nature, with about one-half of the scenarios having the binary "Yes" or "No" responses and the others having either three or four specified categories. In the paper, the scenarios were grouped into content areas such as "fairness of pricing", "importance of incentives", "the perceptions of speculation", "attitudes towards business", and entrepreneurial activities. For the scenarios pertaining to the fairness of pricing, the authors concluded "the reported evidence suggests there is actually little ground that the Soviets are characteristically more hostile toward free-market prices" (p. 390) and that notions of fairness in pricing are very situation-specific.

Whaples (1995) examined how the exposure to economic principles might influence beliefs regarding pricing in the market system. The author administered a survey consisting of six of the scenarios contained in Shiller et al. to 322 students enrolled in 14 sections of an "Introduction to Economics" course. Students in seven sections received the survey (approximately one-half of the students) during the first week of the semester while the other seven sections received the survey at the end of the semester. Whaples not only compared the

pre- and post-course scores with the corresponding scenarios in the Shiller et al. study but also examined the scores by gender. Regarding the pre- and post-scores, Whaples concluded that exposure to economics seemed "to change many students' minds about what is fair, convincing them that market outcomes are equitable" (p. 310). Initially, relative to the male students, female students were considerably less likely to regard the market outcomes as fair. By the end of the semester "female students were still less likely to consider the market outcomes fair, but the gap had narrowed considerably" (p. 310).

THE SURVEY INSTRUMENT AND ASSOCIATED MATERIAL

The survey instrument had 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 consisted of eight scenarios that were used in either the Kahneman et al. study or the Shiller et al. study. The eight scenarios used in this study are presented as Table 1. Six of the eight scenarios pertained directly to a price increase in the market for a good. Some scenarios referenced demand-side effects, some referenced supply-side effects, and one referenced the effect of an increase in a tax. The two non-price scenarios pertained to the effect of a government-administered price ceiling (Scenario 2) and a government quota allotment (Scenario 5).

Three modifications to the scenarios used in the previous studies were enacted for this study. First, the Kahneman et al. study used a total of 18 scenarios, each respondent was asked no more than five scenarios while the Shiller et al. study used a total of 36 scenarios, with each respondent asked 12 scenarios. This study asked each of the respondents the same eight scenarios. Consequently, the sampling design differs from the previous two major studies. Second, the wording of three scenarios was modified slightly from the original studies to reflect societal changes and contextual changes. The three modifications to the original scenarios are the following. Scenario 1 in Table 1 references the price of "a certain product" increasing "after a natural disaster (for example, a tornado, a hurricane, a flood, or a blizzard)" while the original scenario in Kahneman et al. specifically referenced an increase in the price of "snow shovels" after "a large snowstorm." Although a snow shovel is a product to which residents in Toronto and Vancouver could relate, it is not necessarily an appropriate item for all regions in North America. Scenario 7 in Table 1 was also modified slightly. The original question in the Shiller et al. study was "On a holiday, when there is a great demand for flowers, their prices usually go up." Scenario 7 in Table 1 was rewritten to appear as "Before Valentine's Day, florists usually increase the price charged for red roses." A similar change occurred in Scenario 8 in Table 1 as Shiller et al. used "A new railway line makes travel …" but this reference was changed to "A new highway makes travel …". Third, both Kahneman et al. and Shiller et al. reported the results for each scenario as binary responses. As previously noted, Kahneman et al. collapsed the four

categorical responses into two, "Acceptable" and "Unfair", while Shiller used only "Yes" and "No" as the two possible responses. In this study, respondents were asked to respond to the scenarios on the "0% to 100% continuum," with "0%" indicating "Very Unfair" and "100%" indicating "Very Fair." Since very few issues in life related to personal perception are decided in a binary (that is, "black or white") manner, the continuum was deemed the more robust manner in which to gather information and gauge these perceptions.

The survey was administered anonymously during the second week of the Fall 2011 semester in a 100-level (first year) course, Consumer Economics (ECON 110). This course is viewed as a "selective" in one of the topic areas of the University Core Curriculum, as a student can satisfy this requirement by selecting one of five courses listed. This course was desirable to survey for two reasons. First, students enrolled are typically in the first year of university studies, with no previous coursework in economics principles at the university level. Secondly, since the course is a part of the University Core Curriculum, a wide variety of majors will be represented.

THE SURVEY INSTRUMENT AND ASSOCIATED MATERIAL

A total of 181 survey instruments were used in this study (55 from females and 126 from males). The ages of the respondents ranged from 17 to 30, with a mean of 19.6 years and a median of 19 years. In terms of ethnicity, 128 (71%) of the respondents self-identified themselves as Caucasian, while 43 (24%) respondents self-identified themselves as African-American, and seven (4%) more self-identified themselves as Hispanic (or Latino/Latina). In terms of intended major, 84 (46%) of the students indicated they were planning to major in Business and 97 (54%) planning to pursue Non-Business majors (48 in Liberal Arts, 39 in Fine Arts, eight in Education, and two were "Undecided"). For each of the eight statements in the survey, a t-test for difference between means was conducted along gender lines (that is, male and female) and by major field of study (specifically, Business and non-Business).

Examining Differences in Mean Responses by Gender

Whaples observed that, at the start of the economics course, females "were considerably less likely than men to regard the market outcome as fair" (p. 310). Table 2 allows for the examination of the mean responses along gender lines. As previously noted, six of the eight scenarios pertained directly to price changes while the other two involved government involvement in the market. For the six price-related scenarios, all showed males to have a more favorable view of the role of markets. There are two scenarios in which the difference in means is statistically significant at the 6% level. In both Scenarios 3 and 7 males were more accepting of the price increase for the situation portrayed than females. Scenarios 2 and 5 assessed the respondent's view of government involvement in the market. Scenario 2 pertained to the government installing a price ceiling after a natural disaster. Although not statistically significant at the 10% level, females were generally more accepting of such action than males. Scenario 5

pertained to the government restricting gasoline consumption by limiting the amount of gasoline that could be purchased by consumers. Although not statistically significant at the 10% level, males were more accepting of this form of government involvement in the marketplace.

Examining Differences in Mean Responses by Major

Carrithers and Peterson (2006) describe an educational disconnect in the manner in which the role of markets is presented in institutions of higher learning. Although the authors acknowledge the characterization of the two faculty groups may be overly simplistic, the basic premise of their study is that "business and economics faculty focus on the function of markets, the benefits of market economies, and the conduct of business within market economies while A&S faculty focus on flaws and failures of market economies" (p. 373). The authors fear the pedagogical gap will be harmful to students in that if the student hears only one perspective, it "reduces the abilities of our students in their future roles as citizens and leaders" (p. 375).

This study also analyzed the data in terms of major field of study. Table 2 presents the mean responses for the Business/Non-Business students. There are two price-related scenarios in which the difference between the means is statistically significant at the 10% level, both of which were a moderate surprise. The mean response for Business students in Scenario 6 was larger than that for Non-Business majors. At first, this was not what was expected, *a priori*. However, Kahneman et al. concluded that "Judgments of fairness are susceptible to substantial framing effects" (p. 740) and Shiller et al. noted that "notions of fairness are very situation-specific" (p. 389). The initial clause of Scenario 6 frames the major issue with "Suppose the government wishes to reduce the consumption of gasoline". Here, it is not so much the price increase as for the reason for the tax – an attempt to reduce the consumption of gasoline. Scenario 8 referenced raising rents after a new highway has been built. Surprisingly, Non-Business majors thought this was relatively fairer than the Business majors. One of the two non-price scenarios was statistically significant at less than the 1% level. Scenario 5 addressed the government attempt to reduce the consumption of gasoline by limiting the number of gallons purchased by consumers. Business majors thought this initiative was generally "fairer" than did Non-Business majors.

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 between Business and Non-Business majors.

REFERENCES

Carrithers, D.F., & D. Peterson (2006). Conflicting views of markets and economic justice: Implications for student learning. *Journal of Business Ethics*, 69(4), 373-386.

Colander, D. (2003). Integrating sex and drugs into the principles course: Market failures vs. failures of market outcomes. *Journal of Economic Education*, 34(1), 82-91.

Gorman, R.F., & J. B. Kehr (1992). Fairness as a constraint on profit-seeking: Comment. *American Economic Review*, 82(1), 355-358.

Kahneman, D., J. Knetsch & R. Thaler (1986a). Fairness and the assumptions of economics. *Journal of Business*, 59(4, Part 2), S285-S300.

Kahneman, D., J. Knetsch & R. Thaler (1986b). Fairness as a constraint on profit-seeking: Entitlements in the market. *American Economic Review*, 76(4), 728-741.

Prasad, J.N., N. Marlow & R.E. Hattwick (1998). Gender-based differences in perception of a just society. *Journal of Business Ethics*, 17(3), 219-228.

Shiller, R.J., M. Boycko & V. Korobov (1991). Popular attitudes toward free markets: The Soviet Union and the United States compared. *American Economic Review*, 81(3), 385-400.

Whaples, R. (1995). Changes in attitudes among college economics students about the fairness of the market. *Journal of Economic Education*, 26(4), 308-313.

Table 1
THE EIGHT 'FAIRNESS' SCENARIOS

For each of the following questions, please use the following scale:

Very Unfair	**Unfair**	**Moderately Unfair**	**Fair**	**Moderately Fair**	**Very Fair**
0%	**20%**	**40%**	**60%**	**80%**	**100%**

Please indicate your ***perception of the fairness*** of each statement below by writing a number between "0%" and "100%" in the blank to the left of the statement. Please use the numbers between "0" and "100" to reflect the degree to which you agree with the statement. Specifically, if you feel the situation described in the statement is *very unfair* then you should write a number in the blank close to "0" or if you feel the situation described is *generally unfair* then you should write some other number, say "30". Alternatively, if you feel the situation described in the statement was *very fair* then you should write a number close to "100" in the blank or if you feel the situation described was *generally fair* then you should write some other number, say "70".

1.	A store has been selling a certain product for $15. The morning after a natural disaster (for example, a tornado, a hurricane, a flood, or a blizzard) the store raises the price to $30. *To what degree is the increase in this price "fair"?* (Kahneman, et al., #1)
2.	In the situation described above, assume the government establishes a maximum price that limits the price that a business can charge for the product to the pre-disaster price. *To what degree is the government's action to limit the price increase "fair"?* (Shiller, et al., #B3)
3.	A small factory produces tables and sells all that it can make at a price of $200 apiece. Because of reductions in the price of materials, the cost of making each table recently decreased by $20. The factory

	does not change its price of the tables. *To what degree is the decision of the business "fair"?* (Kahneman, et al., #11B)
4.	A small factory produces tables and sells all that it can make at a price of $200 apiece. In fact, the factory cannot produce enough tables to satisfy all the people who want to purchase one. The factory decides to raise the price of the table by $20 even though there was no change in the cost of producing the tables. *To what degree is the increase in this price "fair"?* (Shiller, et al., #B11)
5.	Suppose the government wishes to reduce the consumption of gasoline. The government decides to limit gasoline stations from selling more than five gallons of gasoline to any one person. *To what degree is the government decision to limit the sale of gasoline "fair"?* (Shiller, et al., #C4-1)
6.	Suppose the government wishes to reduce the consumption of gasoline. The government decides to place a major tax on gasoline that will increase the price of gasoline. *To what degree is the government decision to place a tax on gasoline "fair"?* (Shiller, et al., #C4-2)
7.	Before Valentine's Day, florists usually increase the price charged for red roses. *To what degree is this increase in price "fair"?* (Shiller, et al., #B2)
8.	A new highway makes travel between city and summer homes positioned along the highway substantially easier. Accordingly, summer homes along the highway become more desirable and rents on these homes have increased. *To what degree is the increase in the rental price "fair"?* (Shiller, et al., #A9)

Table 2
RESPONSE SUMMARIES AND TESTS OF HYPOTHESES

	Situation/Scenario	Cohort	Characteristic: Mean	St. dev.	$H_1: \mu_x - \mu_y \neq 0$ Pr > | t |
1.	Is it fair for prices to increase after a natural disaster?	**Overall**	**34.867**	**26.018**	
		Females	31.091	24.790	
		Males	36.516	26.464	0.198
		Business	34.167	26.399	
		Non-Bus	35.474	25.807	0.737
2.	Should government limit price increases after a natural disaster?	**Overall**	**61.271**	**24.262**	
		Females	64.546	22.736	
		Males	59.841	24.851	0.231
		Business	59.821	22.461	
		Non-Bus	62.526	25.771	0.456
3.	If the production costs decrease, is it fair if product price does not change?	**Overall**	**61.547**	**22.969**	
		Females	56.636	24.945	
		Males	63.691	21.810	***0.057***
		Business	61.964	21.496	
		Non-Bus	61.186	24.279	0.821
4.	In the presence of a shortage, is it fair for a business to increase price?	**Overall**	**59.337**	**25.212**	
		Females	56.273	26.566	
		Males	60.675	24.587	0.281
		Business	60.000	24.593	
		Non-Bus	58.763	25.850	0.743
5.	To encourage conservation, is it fair for the government to limit the number of gallons of gasoline purchased?	**Overall**	**27.534**	**24.969**	
		Females	23.273	21.714	
		Males	29.135	26.145	0.147
		Business	33.214	26.815	
		Non-Bus	22.278	22.164	***0.003***
6.	To encourage conservation, is it fair for the government to place a tax on gasoline to raise the price?	**Overall**	**26.193**	**22.945**	
		Females	25.818	19.501	
		Males	26.357	22.328	0.885
		Business	30.833	22.722	
		Non-Bus	22.175	22.443	***0.011***
7.	Is it fair to raise the price of flowers before Valentine's Day?	**Overall**	**63.232**	**26.144**	
		Females	55.818	28.460	
		Males	66.468	24.487	***0.011***
		Business	64.167	26.112	
		Non-Bus	62.423	26.281	0.656
8.	Is it fair to raise rents after a new highway is built?	**Overall**	**66.155**	**22.856**	
		Females	62.636	25.219	
		Males	67.691	21.672	0.172
		Business	63.036	24.606	
		Non-Bus	68.856	20.980	***0.088***

ECONOMETRIC TEST OF COST SUBADDITIVITY IN U.S. ELECTRIC INDUSTRY

Deergha R. Adhikari, University of Louisiana at Lafayette
Kishor K. Guru-Gharana, Texas A & M University-Commerce

ABSTRACT

There have been several studies of market power and existence of cost subadditivity in case of U.S cigarette industry and various utility industries. But there is dearth of similar studies in U.S. electric industry. This study attempts to fill that gap. We apply Evans and Heckman's test in the case of cost subadditivity in U.S. electric industry because the electric utility industry in the United States is often cited as an example of a less than perfectly competitive industry. The necessary and sufficient conditions of the test require that the firms chosen for the study have the output at least twice the minimum output observed in the sample. We chose 19 firms that met the conditions. The output quantity for each of the firms was split into the minimum observed quantity and the residual quantity as required by the test. Using a Cobb-Douglas production function the total cost of production for both components of total output (i.e. the minimum quantity and the residual quantity) for each firm were computed and compared with the actual cost of production of the entire quantity by each firm. We found that the sum of the cost of production of the minimum quantity and that of the residual quantity was greater than the cost of production of entire quantity for each firm. Thus, all 19 firms in our sample were found to exhibit cost subadditivity and thereby a natural monopoly.

JEL Classification: L1
Key Words: natural monopoly, cost superadditive, cost subadditive, cost additive

INTRODUCTION

The adherents of deregulation maintain that an increased competition in the markets invariably enhances efficiency in production and distribution. The underlying logic is that efficiency in allocation is achieved as firms after deregulation tend to attain the output level where marginal cost equals price. In so doing, the firms also tend to achieve production efficiency by choosing the input combination that produces a given level of output with a given level of technology at the least costs. But the opponents of deregulation question whether a competitive market always brings about production and allocative efficiency. A situation where the average cost of producing the total demand quantity by a single firm is lower than the average cost of producing the same quantity by two or more firms creates opportunity for a natural monopoly. If that occurs, the cost of producing the total demand quantity would be

minimized by allowing one firm to produce all quantity, that is, by restricting other firms to enter the market.

There have been several studies of market power and existence of cost subadditivity in case of U.S cigarette industry and various utility industries. But there is dearth of similar studies in U.S. electric industry. This is surprising in view of the important position of electric industry in GDP and the lives of millions of people and thousands of other industries affected by their cost/price increase. This study attempts to fill that gap because the electric utility industry in the United States is often cited as an example of a less than perfectly competitive industry. The findings of this study would be important to the Electric Industry companies, its millions of residential and commercial consumers and the Policy makers involved in regulating utility companies.

REVIEW OF UNDERLYING MICROECONOMIC THEORY

A firm, operating in the short run, finds it impossible to vary the quantities of all the inputs it uses in the production, due to its inability to constantly adjust its production capacity to match the ever changing demand for its product. If the capacity of production cannot be adjusted (changed) according to each incremental unit of production then such a situation gives rise to an average cost curve that slopes downward until the capacity is fully exhausted. This situation persists as long as the firm fails to fully adjust its production capacity to every marginal increment in the production. This phenomenon is also referred to as the "economy of scale." To see how the economy of scale gives rise to a downward slopping average cost curve, we differentiate the average cost (AC = C/Y) with respect to the output (Y) as follows:

$$\partial AC/\partial Y = \partial(C/Y) / \partial Y = (Y\partial C/\partial Y - C.\partial Y/\partial Y) / Y^2 = (Y.MC - C) / Y^2 =$$
$$(Y.\ MC/Y - C/Y) / Y = (MC - AC) / Y \qquad (1)$$

where, C is the total cost; and MC is the marginal cost.

As output (Y) can never be negative, this implies that the AC curve slopes downward in the output range where the marginal cost (MC) is smaller than the average cost (AC), a situation called the positive economy of scale. Thus, the AC curve slopes downward so long as a positive economy of scale exists. Conversely, the AC curve slopes upward so long as a negative economy of scale exists, a situation where AC < MC and $\partial AC/\partial Y > 0$.

Within the downward slopping range of the AC curve, it is always cheaper to produce the total demand quantity by one firm than to produce the same quantity by more than one firm. This is illustrated in Figure1below. The AC of producing OQ_2 is lower than the AC of producing OQ_1 by each of the two separate firms where $2OQ_1 = OQ_2$. This situation gives rise to the so called "cost subadditivity" in production. Evans and Heckman (1984) define cost subadditivity as the following. The cost function C (q) is Sub-additive at the output level $\bar{q}$ if and only if

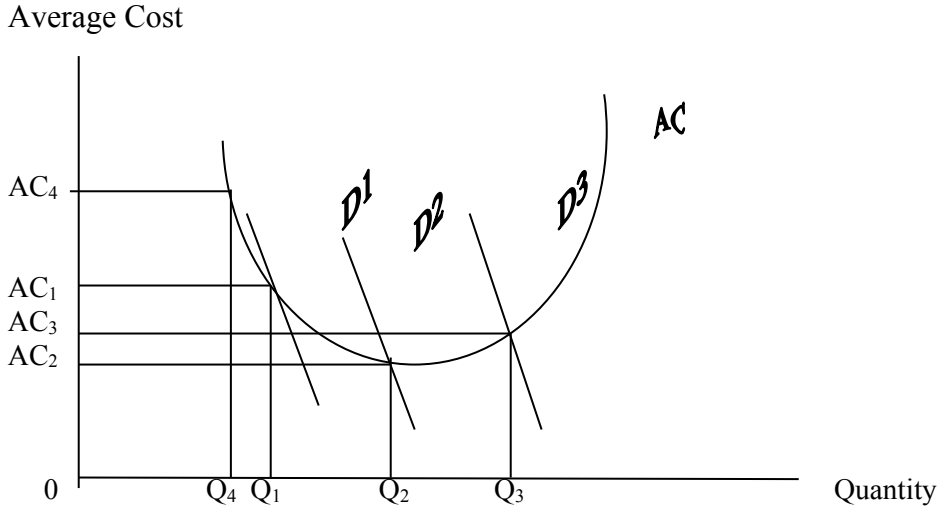
$$C(\bar{q}) < \sum_{i=1}^{n} \sum C(\bar{q}^{i}) \qquad (2)$$

where, $\sum_{i=1}^{n} \bar{q}^{i} / n = \bar{q}$, (3)

and $\bar{q}^{i} \geq 0$ with at least two non-zero vectors of $\bar{q}^{i}$. Here **n** is the number of firms. This very existence of cost subadditivity gives market power to the incumbents by preventing entry to potential entrants and thereby limiting the competition in the market.

As illustrated in Figure 1 below a positive economy of scale implies cost subadditivity. However, as Panzar (1989) argues, positive economies of scale are sufficient but not necessary for the firm's average cost curve to be declining in the single output case. Figure 1 demonstrates that situation.

Figure 1



At OQ_3 level of demand, it is cheaper to produce the total OQ_3 quantity by a single firm (at AC_3) than to produce OQ_2 quantity by one largest firm at the lowest average cost, AC_2, and the residual amount Q_2Q_3 (equal to OQ_4) by a second firm at the average cost, AC_4. Clearly, at OQ_3 level of production, there is a negative economy of scale, but there still exists the cost subadditivity. So, an economy of scale is not required for the existence of cost subadditivity, but the cost subadditivity necessarily exists if there is an economy of scale.

REVIEW OF SELECTED LITERATURE

In Tobacco Case of 1946, the major domestic (U.S.) cigarette manufacturers were accused of operating an illegal cartel (Nicholls, 1949). Although the manufacturers were convicted, but there was a general consensus that the industry behavior was not changed by the verdict. As a result, the aftermath of the case prompted several studies on market conduct and market structure of the cigarette industry.

Sumner (1981) measures the market power of U.S. cigarette industry using a price function. Sullivan (1985) estimates a similar model using a simultaneous equation system approach over the year 1955-82. A study by Adhikari (2004) measures the market power of U.S. cigarette industry using revenue elasticity approach. Furthermore, one can find several studies done on cost subadditivity in various public or private utilities companies. Studies in this group include those by Bitzan (2003), Sueyoshi (1996), Cubukcu et al (2008), Jamasb et al (2008), Everett (2008), Yudong et al (2008), Currier et al (2008), Wills-Johnson (2008), Won (2007), Fung et al (2007), Becker et al (2006), Kwoka (2006), Chang et al (2006), Ramos-Real & Javier (2005), and Gordon et al (2003). These studies cover communication and transportation except for Ramos-Real & Javier (2005). There is, however, lack of empirical study on the tests for cost subadditivity in U.S. electric industry. Most of the studies mentioned above apply a translog function for the estimation of the cost function and for the determination of cost subadditivity. However, these studies don't impose a necessary or a sufficient condition to test their hypothesis. Our study will apply Evans and Heckman's test (1984) for the test of cost subadditivity and will test the hypothesis by imposing both necessary and sufficient condition.

We will present the model for the study in section 3. In section 4 we will explain the data and the methodology of the study. The empirical findings will be presented in section 5 and will summary of our results in section 6.

THE MODEL

Baumol et al. (1982) have recommended separate tests for necessary and sufficient conditions for cost subadditivity. Because, doing so will allow the researcher to reject the hypothesis of cost subadditivity if the necessary condition fails to be satisfied, and to accept it if the sufficient condition is met. However, the problem with this testing procedure, in a single product case, is that the test becomes inconclusive if the acceptance of the necessary condition occurs together with the rejection of the sufficient condition. Therefore, this study applies Evans and Heckman's test for the test of cost subadditivity. They derive the test as following.

Since an industry can be split into two or more firms in an infinite number of ways, a global test for cost subadditivity is extremely difficult. Owing to this problem, Evans and Heckman have developed a local test for cost subadditivity. By employing certain restrictions, as determined by observed data points, they have narrowed down the area over which the test could be applied. The region confined within these restrictions is called the "admissible region." For the sake of simplicity, they assume that there only exist two firms in the industry, and so, n = 2. Denoting the first hypothetical firm by A, and the second by B the total output can, then, be expressed as $q = q^A + q^B$. The cost of production of the total quantity, q, by the two firms is $C^A + C^B$, whereas the cost of producing the whole quantity, q, by a single firm is C. If $C < C^A + C^B$ for all two-firm configurations, then the cost function is subadditive at q, over an admissible region. They specify two constraints that define the admissible region.

The first constraint requires that no hypothetical firm be permitted to produce less of either of the two outputs than the output of the firms for which there is data. Suppose q_m is the vector of minimum output such that $q_m = (\text{min. } q_{1t} , \text{min.} q_{2t}) = (q_{1m} , q_{2m})$ where min. q_{it} is the minimum quantity of *i*th output. Suppose firm A and B produce as following:

$q_t^A = (\phi\, q_{1t}^* + q_{1m},\ w\, q_{2t}^* + q_{2m})$; and (4)
$q_t^B = [(1-\phi)\, q_{1t}^* + q_{1m},\ (1-w)\, q_{2t}^* + q_{2m}]$, (5)

where, q_{it}^* is the incremental quantity and q_{im} is the minimum quantity of ith output respectively. Then the industry production of output 1 and 2 can be expressed as

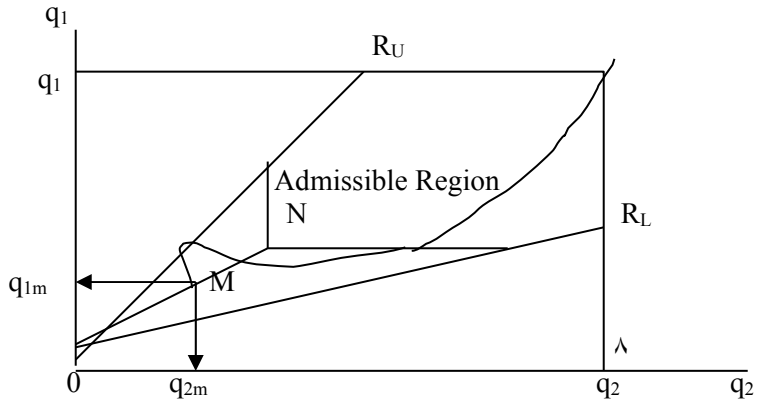
$\bar{q}_{1t} = q_{1t}^* + 2q_{1m}$ (6)
$\bar{q}_{2t} = q_{2t}^* + 2q_{2m}$ (7)

Thus the first constraint requires that the test be based on the firms for which the output of each of the two products is at least twice the output level in the sample. This constraint also holds in one product case. The second constraint requires that both firms A and B produce q_1 and q_2 in a ratio within the range of the ratios observed in the data. This implies the following inequalities.

$R_L < (\phi q_{1t}^* + q_{1m}) / (w\, q_{2t}^* + q_{2m}) < R_U$ (8)
$R_L < [(1-\phi)\, q_{1t}^* + q_{1m}] / [(1-w)\, q_{2t}^* + q_{2m}] < R_U$, (9)

where R_L is min.(q_{1t}/q_{2t}) and R_U is max.(q_{1t}/q_{2t}). The admissible region can be shown as following.

Figure 2



At point M, both outputs are at their observed minimum levels, whereas, at point N, both output quantities are double their observed minimum levels. Therefore, any point to the right of point N will satisfy the first constraint. Since R_U and R_L are maximum ratio of output 1 to output 2 in observed data, any point to the right of OR_U line but to the left of OR_L line satisfies the second constraint. Since q_1 and q_2 are the maximum levels of outputs observed in the data, the admissible region satisfies all the constraints. So, the test of sub-addativity has to be limited within the admissible region. In one product case the second constraint reduces to the following inequalities

$$\text{min. } q_t \leq \phi\, q_t^{*} + q_m = \bar{q}_t^{A} \leq \text{max. } q_t \qquad (10)$$

$$\text{min. } q_t \leq (1-\phi)\, q_t^{*} + q_m = q_t^{B} \leq \text{max. } q_t \qquad (11)$$

It means that none of the hypothetical firms be permitted to produce lower than the observed minimum quantity and higher than the observed maximum quantity. In one product case the second constraint is satisfied by all the observations. Therefore, in one product case, the observations considered for the test of sub-addativity have to satisfy the first constraint only. As such, only those observations can be taken for the test, which have output quantity twice as much as the minimum observed quantity.

Let $\bar{C}(\bar{q}_t^{A})$, $\bar{C}(q_t^{B})$ and $\bar{C}(\bar{q}_t^{A} + \bar{q}_t^{B})$ be the cost of producing $\bar{q}_t^{A}$ and $\bar{q}_t^{B}$ by firm A and firm B and the cost of producing $\bar{q}_t^{A} + \bar{q}_t^{B}$ by a single firm respectively. Then the degree of cost subadditivity is measured by:

$$\text{SUB} = [\bar{C}(\bar{q}_t^{A} + \bar{q}_t^{B}) - \bar{C}(\bar{q}_t^{A}) - \bar{C}(\bar{q}_t^{B})] / \bar{C}(\bar{q}_t^{A} + \bar{q}_t^{B}) \qquad (12)$$

If SUB is less than zero the cost function is Sub-additive; if it is zero the cost function is additive; and if it is greater than zero, then the cost function is super-additive.

THE DATA AND METHODOLOGY

The data on all the variables (e.g. input costs and output) for the U.S. electric industry have been obtained from the United Nations Industrial Development Organization (UNIDO) website: http://www.unido.org. Only those firms have been chosen for the study for which the output is at least double of the minimum quantity observed in the sample in order to satisfy the constraints defined in our model. The relevant data set on all the firms is given in Appendix A.

The costs of producing total output have been estimated using the Cobb-Douglas cost function as shown in the appendix-A. To test for cost subadditivity, we need to split each firm's output into two or more parts. A firm's output quantity can be split into two parts in infinite number of ways without violating the constraint. However, to economize on time, output in each observation has been split into the minimum observed quantity, which is 0.248 million kilowatt hours, and the residual quantity. Then using the estimated cost function, the cost for each of the two components has been estimated for each firm. Based on the above estimates, the degrees of subadditivity have been estimated for each of the admissible firm using equation (12). Values less than zero for the variable SUB imply cost subadditivity; zero implies cost additivity and values greater than zero for the variable SUB_t imply super-additivity.

EMPIRICAL FINDINGS

Based on the data on 19 firms, we estimated the following Cobb-Douglas cost function:

LNCOST = -0.329753698+0.774837428LNPL–0.421208228LNPK + 0.840977191 LNY
(-0.7786451) (3.057291524) (-1.58613098) (17.36975228)
R^2 = 0.9583, F-statistic = 114.8199, p-value associated with the F-value = 0.000

The value LNCOST is the log of the long run total cost for the generation and transmission of electric power expressed in millions of dollars, LNPL is the log of the average annual payment per worker expressed in thousands of dollars, LNPK is the log of the estimated user cost of capital expressed in thousands of dollars, and LNY is the log of the total generation and transmission of electric power expressed in millions of kilowatt-hours. The data on these variables are given in the appendix. The figures in the parentheses are the associated t-values. The coefficients associated with LNPL and LNY are significant at 1 percent level whereas that associated with the variable LNPK is not significant even at 10 percent level. A high R^2 value indicates that the model fits the data well and the p-value associated with the F-statistic indicates that the Coefficient of Determination is highly significant. Therefore, we use this model to estimate the total long-run cost of producing the total quantity as well as the cost of producing both the minimum quantity ((0.248 millions of kilowatt-hours), and the residual quantity for each firm. We, compute the average cost of production of the total quantity, the minimum quantity, and the residual quantity as the following:

Log of average cost of production = LNCOST – LNY (13)

The average cost of production for the total and for each quantity for each firm is, then, estimated by taking the exponent of the log of the average cost of production, which is shown in Appendix-B. The average cost of producing the minimum quantity and that of the residual quantity are added together. Finally, the sum was subtracted from the average cost of producing the total quantity for each firm. The result is the measure of cost-additivity (the result is shown in column COSTADD in Table 1 below. A negative entry indicates that the sum of the average cost of producing the minimum quantity and that of the residual quantity is greater than the average cost of producing the whole quantity, exhibiting thereby the cost subadditivity. The results in Table 1 show that the average cost of production of each firm is sub-additive.

Table 1

Subadditivity Estimates for 19 Private U.S. Electric Utility Firms

Firm	ECOST	ECOST1	ECOST2	COSTADD	SUB
1	2.284075	2.656892	0.199748	-0.57256485	Sub-additive
2	2.33695	2.795377	0.171976	-0.6304033	Sub-additive
3	2.508023	2.823898	0.252225	-0.56810021	Sub-additive
4	2.114484	2.755656	0.196873	-0.83804574	Sub-additive
5	2.362858	2.937729	0.215758	-0.79062906	Sub-additive
6	2.54975	2.796534	0.26923	-0.5160153	Sub-additive
7	2.230965	2.819163	0.190253	-0.77845152	Sub-additive
8	2.281219	2.91201	0.175835	-0.80662636	Sub-additive
9	2.208226	2.609671	0.256217	-0.65766195	Sub-additive
10	2.4956	2.94253	0.198923	-0.64585272	Sub-additive
11	2.238234	2.790404	0.165291	-0.71746078	Sub-additive
12	2.052351	2.893871	0.149038	-0.99055829	Sub-additive
13	2.081873	2.541041	0.141839	-0.60100668	Sub-additive
14	1.872957	2.601372	0.145678	-0.8740938	Sub-additive
15	2.03678	2.663727	0.231515	-0.85846204	Sub-additive
16	1.920776	2.665863	0.163748	-0.90883421	Sub-additive
17	2.256829	2.816801	0.177496	-0.73746772	Sub-additive
18	2.190043	2.788608	0.164746	-0.76331088	Sub-additive
19	2.561288	2.960528	0.21376	-0.61300022	Sub-additive

ECOST = Exponent of **LNCOSTY** = Long-run average cost for the entire quantity for the firm

ECOST1 = Exponent of **LNCOST1Y** = Long-run average cost for the minimum quantity for the firm

ECOST2 = Exponent of **LNCOST2Y** = Long-run average cost for the residual quantity for the firm

COSTADD = **ECOST - (ECOST1 + ECOST2)** = Long-run average cost for the entire quantity minus sum of the average cost for the minimum quantity and the average cost for the residual quantity

SUB = A measure of cost-subadditivity (If **COSTADD** is less than zero, the firms average cost is sub-additive).

SUMMARY OF RESULTS

If the average cost of producing the whole demand quantity by a single firm is lower than that of producing the same quantity by two or more firms combined, then such situation gives rise to a natural monopoly. In this situation the cost of producing the whole demand quantity is minimized by allowing one firm to produce all quantity. The electric utility industry in the United States is often cited as an example of a natural monopoly. Our study applies Evans and Heckman's test for the test of cost subadditivity on U.S. electric industry. The necessary and sufficient conditions of the test require that the firms chosen for the study have the output at least twice the minimum output observed in the sample. We chose 19 firms that met the conditions. Then the output quantity for each of the firms was split into the minimum observed quantity and the residual quantity as required by the test. Using a Cobb-Douglas production function the total cost of production for each of the quantities (i.e. the minimum quantity and the residual quantity) for each of the firms were computed and compared with the actual cost of production of the entire quantity by each firm. We found that the sum of the cost of production of the minimum quantity and that of the residual quantity was greater than the cost of production of entire quantity for each firm. Thus, each of the firms in our sample was found to exhibit cost subadditivity and thereby a natural monopoly. In simple words it is more cost effective to let the existing industries grow to fulfill the growing demand compared to entry of new companies. This finding is important for millions of consumers, the existing electric companies and the policy makers because it provides strong basis for regulating entries into this industry.

REFERENCES

Adhikari, Deergha (2004). Measuring Market Power of U.S. Cigarette Industry. *Applied Economics Letters*, 11(15), 957-59.

Baumol, W. J.; J. C. Panzar and Robert D. Willig (1982). Contestable Markets and the Theory of Industry Structure, San Diego: Harcourt, Brace, Jovanvich.

Becker, Klaus G., Von Ende, Eleanor T. (2006). Sub-Game Perfect Equilibria in Natural Monopoly Markets. *Atlantic Economic Journal*, 34(3), 365-366.

Berndt, Ernst R. (1991). The Practice of Econometrics: Classic and Contemporary.Addison-Wesley Publishing Company, Inc.

Bitzan, John D. (2003). Railroad Costs and Competition. *Journal of Transport Economics & Policy,* 37(2), 201-225.

Chang, Hsihui, Mashruwala, Raj (2006). Was the bell system a natural monopoly***?*** An application of data envelopment analysis. *Annals of Operations Research*, 145(1), 251-263.

Cubukcu, K. Mert, Guldmann, Jean-Michel (2008). Geography and the costs of local telephone networks. *Annals of Regional Science*, 42(4), 821-842.

Currier, Kevin M., Jackson, Brian K (2008). A Pedagogical Note on the Superiority of Price-Cap Regulation to Rate-of –Return Regulation. *Journal of Economic Education*, 39(3), 261-268.

Evans, David S. and Heckman, James J. (1984). A Test for Subadditivity of the Cost Function with an Application to the Bell System. *American Economic Review*, 74, 615-623.

Everett, Sophia (2008). Public sector deregulation: a paradigm in conflict.*Maritime Policy & Management*, 35(4), 341-351.

Fung, K. K., Kolar, Sri Harsha, Karnam, Pavan (2007) " Regulating Natural Monopoly—Flash Animation," Journal of Economic Education; 38(2), 255-255.

Gordon, D.V., Gunsch, K., Pawluk, C.V. (2003). A natural monopoly in natural gas Transmission. *Energy Economics*, 25(5), 473.

Gujarati, Damodar N. (1995), Basic Econometrics, McGraw Hill, Inc.

Jamasb, Tooraj, Pollitt, Michael (2008). Security of Supply and Regulation of Energy Networks. *Energy Policy*, 36(12), 4584-4589.

Kwoka, John E. (2006) "The Role of Competition in Natural Monopoly: Costs, Public Ownership, and Regulation. *Review of Industrial Organization*, 29(1/2), 127-147.

Kristin Boose (1991). An Analysis of Subadditivity of Costs as Applied to the Deutsche Bundespost. (Unpublished).

Nicholls, William H. (May 1949). The Tobacco Case of 1946. *American Economic Review: Papers and Proc*.39(3), 284-296.

Panzar, John C., 1989. Technological determinants of firm and industry structure. In Handbook of Industrial Organization, R. Schmalensee & R. Willig (ed.), ed. 1, vol.1, ch. 1, 3-59, Elsevier.

Ramos-Real, Francisco Javier (2005). Cost functions and the electric utility industry. A contribution to the debate on deregulation. *Energy Policy*, 33(1), 69-87.

Sullivan, Daniel (1985). Testing hypothesis About Firm Behavior in the Cigarette Industry. *Journal of Political Economy*, 93(1-3), 586-598.

Sumner, Daniel A. (1981). Measurement of Monopoly behavior: An Application to the Cigarette Industry. *Journal of Political Economy*, 89(4), 1010-1019.

Sueyoshi, Toshiyuki (1996). Divestiture of Nippon Telegraph and Telephone. *Management Science*, 42(9), 1326-1351.

Won, Gu-Hwan (2007). Electric power industry restructuring and ROE: The case of Korea Electric Power Corporation. *Energy Policy*, 35(10), 5080-5090.

Wills-Johnson, Nick (2008). Separability and Subadditivity in Australian Railways.*Economic Record*, 84(264), 95-108.

Yudong Qi, Xuexin Liu, Zhili Wang (2008). Governance Model Reform in Natural Monopoly Industries. *Chinese Economy,* 4(4), 5-22.

Appendix-A

Sample Data for 20 Private U.S. Electric Utility Firms

Firm	C	Q	r	w
1	30.8923	4.612	0.06903	8.5368
2	58.5825	8.297	0.06903	9.9282
3	15.1205	1.82	0.06754	10.1116
4	32.8014	5.849	0.07919	10.2522
5	22.7768	3.145	0.06481	11.1194
6	11.9176	1.381	0.06598	9.6992
7	34.4028	5.422	0.06754	10.0613
8	47.5209	7.115	0.06565	10.9087
9	18.9136	3.052	0.10555	10.1954
10	36.0902	4.394	0.06572	11.2585
11	62.0032	9.699	0.06903	9.8758
12	74.7206	14.271	0.06789	10.9051
13	96.0053	17.743	0.06903	7.4775
14	63.4357	14.956	0.06572	7.8062
15	15.9901	3.108	0.07919	9.2689
16	42.3249	9.416	0.06565	8.3906
17	44.6781	6.857	0.06565	9.8826
18	59.252	9.745	0.0686	9.8235
19	38.7337	4.442	0.08206	12.9352

C= Total long-run cost of generation and transmission of electric power, expressed in millions of dollars

Q= Total generation and transmission of electric power, expressed in millions of kilowatt-hours.

r= Estimated user cost of capital, $r=q_k(i+\delta)$, where q_k is the unit acquisition cost of the capital stock, I is the real rate of interest and δ is the rate of depreciation.

w= Average annual payment per worker, expressed in thousands of dollars.

Source: United Nations Industrial Development Organization (UNIDO) website http://www.unido.org.

Appendix-B

Cost Estimates For 19 Private U.S. Electric Utility Firms

Firm	C	Q	r	w	LNCOST1	LNCOST2	LNCOSTY	LNCOST1Y	LNCOST2Y
1	30.8923	4.612	0.06903	8.5368	0.371609	-0.97081	0.8259609	0.9771569	-1.6106996
2	58.5825	8.297	0.06903	9.9282	0.422419	-0.85466	0.8488468	1.0279671	-1.760399
3	15.1205	1.82	0.06754	10.112	0.43257	-1.18098	0.9194948	1.0381183	-1.3774345
4	32.8014	5.849	0.07919	10.252	0.408107	-0.87693	0.7488108	1.0136556	-1.6251949
5	22.7768	3.145	0.06481	11.119	0.472089	-1.07165	0.8598721	1.0776369	-1.533597
6	11.9176	1.381	0.06598	9.6992	0.422833	-1.25796	0.9359951	1.0283809	-1.3121878
7	34.4028	5.422	0.06754	10.061	0.430892	-0.94557	0.8024343	1.0364402	-1.6593995
8	47.5209	7.115	0.06565	10.909	0.463295	-0.90144	0.8247098	1.0688436	-1.7382101
9	18.9136	3.052	0.10555	10.195	0.353676	-0.91395	0.7921897	0.9592242	-1.3617295
10	36.0902	4.394	0.06572	11.259	0.473721	-0.99721	0.9145292	1.0792697	-1.614838
11	62.0032	9.699	0.06903	9.8758	0.420638	-0.82457	0.8056871	1.0261864	-1.8000491
12	74.7206	14.27	0.06789	10.905	0.457047	-0.75671	0.7189859	1.0625950	-1.9035519
13	96.0053	17.74	0.06903	7.4775	0.327025	-0.71015	0.7332682	0.9325738	-1.953061
14	63.4357	14.96	0.06572	7.8062	0.350491	-0.7588	0.6275183	0.9560390	-1.9263536
15	15.9901	3.108	0.07919	9.2689	0.374178	-1.00675	0.7113702	0.9797264	-1.4631116
16	42.3249	9.416	0.06565	8.3906	0.374979	-0.84715	0.6527295	0.9805277	-1.8094264
17	44.6781	6.857	0.06565	9.8826	0.430053	-0.90867	0.8139605	1.0356017	-1.7288088
18	59.252	9.745	0.0686	9.8235	0.419994	-0.82577	0.7839212	1.0255426	-1.8033527
19	38.7337	4.442	0.08206	12.935	0.479819	-0.92027	0.9405104	1.0853677	-1.5428999

LNCOST1 = Log of long-run total cost estimate for the minimum quantity for the firm

LNCOST2 = Log of long-run total cost estimate for the residual quantity for the firm

LNCOSTY = Log of long-run average cost for the entire quantity for the firm

LNCOST1Y = Log of long-run average cost estimate for the minimum quantity for the firm

LNCOST2Y = Log of long-run average cost estimate for the residual quantity for the firm

THE DEBT INDEX AND ITS RELATION TO ECONOMIC ACTIVITY: AN EXTENSION

John J. Bethune, Barton College

INTRODUCTION

During the 1970s and 80s the concept of a "misery index" was used as a proxy to describe how well, or poorly, the macro economy was performing. In its simplest form the misery index was calculated by adding the rate of inflation to the rate of unemployment, thus a higher index indicated an economy preforming poorly.

In the past two decades, with inflation seemingly under control and, until recently, a modest level of unemployment, the misery index has not been the subject of policy discussions or political discourse. Rather, concern about the national debt and soaring budget deficits seems to be the focus of those who worry about our economic future.

With expanding national indebtedness and seemingly endless deficit spending the world's economies appear to face different issues that move beyond inflation, unemployment, and sluggish growth rates. While there are clearly empirical relationships for these variables to debt and deficit levels, until recently we did not have an index that shows explicitly how debt affects economic activity.

This paper expands on a previous publication that combines debt to GDP and deficit to federal spending ratios to develop a "debt index" for several national economies. While the earlier effort used measures of the debt index to compare with various macroeconomic variables, this work will compare the movement of the debt indices through time with the movement of macroeconomic variables across 14 countries. Given the characteristics of the data, this approach is more appropriate than my previous effort. Also, this paper includes regression analysis to gauge the explanatory power of the relationships. Given the characteristics of the data (see below) this is appropriate for the relative change data used in this analysis, but would have been inappropriate for use in the prior data set.

The Debt Index and Prior Research

There are both short run and long run issues involving the debt problem in the United States and elsewhere. In the short run, the deficit represents a problem for policy makers while in the long term, the national debt is an issue that must be addressed.

To construct a "debt index" I combine the value of the annual federal budget deficit divided by federal government spending with the national debt divided by nominal GDP.
Put simply:

Deficit/Spending + Debt/GDP = Debt Index

This combines the temporal aspects of our short and long term debt problems into one measure.

In a recent article I use correlation coefficients to show how the debt index is associated with private investment and the rate of unemployment for 15 industrialized countries. [Bethune,

2013] For the United States, I show how the debt index is more closely correlated, in most instances, with private investment and unemployment than any of the component parts. Other topics are addressed as well.

Another recent article finds a relationship between budget deficits and economic growth, which reinforces my more comprehensive study. [Cebula, 2013] Focusing only on budget deficits, Cebula finds that "the higher the budget deficit (expressed as a percent of GDP), the lower the percentage growth rate of real per capital GDP." [p.86]

While there are other studies examining the relationship between deficits, debt and various macroeconomic variables, none use the temporal index I developed in the previous article, thus an extensive literature review is not possible. These prior studies only focus on how deficits affect economic growth and do not address the issue of overall debt. The basic content of this research approach is unique.

Extensions of the Relationships

While the earlier work just used the values of the variables as they moved through time, it is possible to make additional meaningful statistical comparisons by examining the data in the form of percentage change from one time period to the next. Using the Pearson product movement correlation coefficient, Table I presents the relationship between the debt index and private investment for 14 industrialized countries. [The data set for Iceland, included in the previous study, was corrupted, and I did not take the time to reconstruct it, given that it added little overall relevance.] Unless noted otherwise, all coefficients are significant at the one percent level.

Table I
Debt Index Correlations with Private Investment (Percentage Change)

USA	-.672
Greece	-.501
Italy	-.686
Japan	-.782
Sweden	-.787
UK	-.723
Germany	-.392 **
Australia	-.365**
New Zealand	-.467*
Canada	-.818
France	.653
Ireland	-.533
Spain	-.550
Portugal	-.665

*Significant at the 05 percent level.
**Significant at the 10 percent level.
@ Not significant at the 10 percent level.

`

In the initial study, 13 of the 15 countries demonstrated significant relationships between the two variables. Using the percentage change in the variables, all the countries here demonstrate some significant correlation. In most cases, however, the correlation is somewhat weaker. For example, the data for the USA correlated at -.831, but using the percent change in the variables results in a -.672 coefficient. Both remained significant at the one percent confidence interval.

Table II presents the relationship between the debt index and unemployment. In the previous study the UK, Germany and France showed no significant relationship. Using this method, only Germany continues to exhibit a weak and insignificant relationship. For Greece, the sign change was reversed, indicating unemployment and debt are not related in a manner similar to the rest of the countries.

For the USA, the relationship strengthened from .479 to .802, indicating a strong positive association with debt and the rate of unemployment.

Table II
Debt Index Correlations with the Unemployment Rate (Percentage Change)

USA	.802
Greece	-.174
Italy	.601
Japan	.804
Sweden	.608
UK	.661
Germany	.183@
Australia	.623
New Zealand	.533
Canada	.845
France	.505*
Ireland	.448
Spain	.843
Portugal	.481*

*Significant at the 05 percent level.
**Significant at the 10 percent level.
@ Not significant at the 10 percent level.

Table III presents the relationship between total spending in the private sector and the debt index. This relationship was not examined in the prior research. As indicated, except for Germany, there is generally a strong negative association between the debt index and private sector spending.

Table III **Debt Index Correlations with Private Sector Spending (Percentage Change)**	
USA	-.713
Greece	-.660
Italy	-.677
Japan	-.737
Sweden	-.740
UK	-.836
Germany	-.183@
Australia	-.439*
New Zealand	-.495
Canada	-.818
France	.712
Ireland	-.459
Spain	-.596
Portugal	-.604

*Significant at the 05 percent level.
**Significant at the 10 percent level.
@ Not significant at the 10 percent level.

Another relationship not presented in the previous article is that between the debt index and nominal GDP. Table IV shows these correlation coefficients. Germany, Australia and Spain do not exhibit any significant relationship, but the rest of the remaining countries show a significant negative relationship. Higher increases in the debt index coefficients are associated with slower or negative rates of growth in GDP.

Table IV **Debt Index Correlations with GDP (Percentage Change)**	
USA	-.527
Greece	-.815
Italy	-.329**
Japan	-.614
Sweden	-.760
UK	-.693
Germany	-.198@
Australia	-.293@
New Zealand	-.339**
Canada	-.850
France	.377*
Ireland	-.515
Spain	-.209@
Portugal	-.536

*Significant at the 05 percent level.
**Significant at the 10 percent level.
@ Not significant at the 10 percent level.

Overall, for most countries in most cases, increasing debt indexes are associated with decreasing (or negative) rates of growth for private investment, private spending levels and nominal GDP. Also the debt index movements are positively and significantly associated with the changes in the rate of unemployment.

The Debt Index and Explanatory Power

While correlation coefficients can show how closely two variables are associated this does not necessarily demonstrate the causal relationship between the variables. During the course of the prior research I attempted some simple regressions to test the explanatory power of the correlated relationships. These did not offer any additional insight.

In retrospect we should not expect simple regression models to offer much in the way of explanatory power. These models assume that Y is a linear function of X and are appropriate "when X and Y are stationary time series or cross-sectional (non-time-series) variables, and a scatter plot of Y versus X suggests a significant linear relationship." [Duke Website] It is much more probable that the percentage change in Y is a linear function of the percentage change in X, in which case a relative change model would be preferable. This model is also appropriate "when X and Y are nonstationary time series with nonlinear trends and/or heteroscedasticity--e.g., series with inflationary or compound growth . . ." [Duke Website] which would appear to be the case here. The variables do contain inflationary growth and heteroscedasticity is likely.

In the following four tables I present the results from a regression model that uses the percent change in the debt index as the independent variable to explain the percentage change in the variables addressed in the four previous tables.

Table V
Adjusted R-square Where Private Investment is the Dependent Variable and the Debt Index is the Explanatory Variable (Percentage Change)

USA	43.4 percent
Greece	0
Italy	44.6 percent
Japan	59.9 percent
Sweden	59.8 percent
UK	50.6 percent
Germany	10.9 percent
Australia	9.2 percent
New Zealand	18.7 percent
Canada	64.7 percent
France	40.8 percent
Ireland	26.0 percent
Spain	28.0 percent
Portugal	41.4 percent

Table VI **Adjusted R-square Where the Unemployment Rate is the Dependent Variable and the Debt Index is the Explanatory Variable (Percentage Change)**	
USA	63.2 percent
Greece	12.8 percent
Italy	33.2 percent
Japan	63.5 percent
Sweden	33.3 percent
UK	41.9 percent
Germany	0
Australia	35.9 percent
New Zealand	25.6 percent
Canada	69.5 percent
France	23.1 percent
Ireland	17.4 percent
Spain	70.1 percent
Portugal	19.3 percent

Table VII **Adjusted R-square Where Private Sector Spending is the Dependent Variable and the Debt Index is the Explanatory Variable (Percentage Change)**	
USA	49.2 percent
Greece	0
Italy	43.3 percent
Japan	52.7 percent
Sweden	52.2 percent
UK	69.0 percent
Germany	0
Australia	15.4 percent
New Zealand	21.4 percent
Canada	64.6 percent
France	49.1 percent
Ireland	18.4 percent
Spain	33.4 percent
Portugal	33.3 percent

Table VIII **Adjusted R-square Where Nominal GDP is the Dependent Variable and the Debt Index is the Explanatory Variable (Percentage Change)**	
USA	25.4 percent
Greece	3.1 percent
Italy	6.8 percent
Japan	35.6 percent
Sweden	55.3 percent
UK	46.2 percent
Germany	0
Australia	4.2 percent
New Zealand	8.0 percent
Canada	70.4 percent
France	11.4 percent
Ireland	24.0 percent
Spain	1.2 percent
Portugal	25.2 percent

With respect to private investment, the debt index has considerable explanatory power (greater than 40 percent) for the USA, Italy, Japan, Sweden, the UK, Canada, France, and Portugal. The other countries exhibit a moderate to weak causal relationship.

For the unemployment rate, the debt index has considerable explanatory power for the US, Japan, the UK, and Spain. The other countries exhibit a moderate to weak causal relationship.

For private sector spending, the debt index has considerable explanatory power for the USA, Italy, Japan, Sweden, the UK, Canada and France. The other countries exhibit a moderate to weak causal relationship.

Finally, for nominal GDP, only the UK, Sweden and Canada show an explanatory power of greater than 40 percent. Moderate (adjusted R-squares between 20 and 39.9 percent) are present for the USA, Japan, Ireland and Portugal.

The growth (or lack of) in GDP is often considered as having considerable explanatory power with respect to private investment, private sector spending and the unemployment rate. Using USA data I ran a simple regression where nominal GDP was used as the explanatory variable.

The percentage change in nominal GDP does outperform the debt index in explaining the percent change in private investment (56.1 percent v. 43.4 percent) and the percent change in the private sector (85.6 percent v. 43.4 percent). The power of nominal GDP to explain these variables was generally stronger in all countries where the debt index had considerable explanatory power as well.

However, the debt index did outperform nominal GDP with respect to unemployment (63.2 percent v. 37.5 percent). This was true in all the countries that showed considerable explanatory power for the debt index: Japan (63.5 percent v. 21.4 percent), the UK (41.9 percent v. 33.7 percent), and Spain (70.1 percent v. 10.7 percent).

The Debt Index and Forecasting

The International Monetary Fund forecasts future data through the year 2016. Table IX shows the percentage change from one year to the next for the USA in these annual forecasts. [The debt index forecast is calculated based on the IMF forecasted data of the component parts.] It is predicted that the debt index will fall in 2013 and then gradually rise each year through 2016.

Also predicted is relatively robust growth in private investment, the private sector, and GDP, while unemployment declines significantly.

Table IX
IMF Annual Forecasts for the USA (Percent Change)

Year	PC Debt Index	PC Private Investment	PC Private Sector	PC GDP	PC Unemployment
2013	-01	7.2	4.1	3.2	-5.6
2014	00	9.0	3.5	4.0	-8.6
2015	1.6	9.1	3.7	4.7	-11.0
2016	2.1	8.2	3.6	4.9	-12.7

If these predictions are accurate the relationship between the debt index and these variables will change significantly over the next four years. Table X compares the past correlation coefficients (taken from Tables I through IV) of the debt index and the relevant variable with the forecasted correlations.

Table X
Correlation Comparisons

Private Investment		Unemployment		Private Sector		GDP	
Previous	Predicted	Previous	Predicted	Previous	Predicted	Previous	Predicted
-.672	.272@	.802	-.595@	-.713	-.030@	-.527	.131@

@ not significant at the 10 percent level

While, since 1980, there have been strong negative correlations (significant at the one percent level) between the debt index and private investment, the private sector, and GDP, the forecasted associations are insignificant and change signs in two instances. The same holds true for the predicted association between the debt index and unemployment (a sign change and no significance).

Table XI
Adjusted R-square Comparisons with the Debt as the Explanatory Variable

Private Investment		Unemployment		Private Sector		GDP	
Previous	Predicted	Previous	Predicted	Previous	Predicted	Previous	Predicted
43.4 percent	0.0 percent	63.2 percent	13.9 percent	49.2 percent	0.0 percent	25.4 percent	44.9 percent**

** Indicates a positive relationship between the two variables, opposite sign of the previous relationship.

Table XI shows the comparison between previously calculated adjusted R-squares (from Tables V through VIII) and those predicted using the IMF data.

The relationships between the IMF's predictions and the debt index are clearly at odds with the previous relationships held between these variables. The IMF is forecasting a moderately rising debt index concurrent with robust growth in private investment, the private sector, and GDP. It is also suggesting accelerating decreases in the unemployment rate.

The results of the research presented here suggest that, if the debt index continues to rise, much more anemic growth will occur in private investment, the private sector, and GDP. We would also expect very little progress towards reducing the unemployment rate. Further, if the IMF forecasts that call for moderate growth in the debt index are overly optimistic then we would expect an even worse performance from these macro-variables. Time will tell.

SUMMARY AND CONCLUSIONS

As noted in my earlier work, policy makers and commentators do not currently have a useful index to track or use to show how debt affects current and future economic activity. The debt index offers such a tool and combines both short term and long term considerations.

Using correlation coefficients and relative change regression analysis it can be shown that, for a variety of countries, the debt index is significantly and adversely related to such variables as private investment, unemployment, private spending, and nominal GDP. Correcting both short term problems (the deficit to government spending ratio) and longer term problems (the national debt to GDP ratio) might well be the key to increasing levels of private sector investment and spending and, thus, increasing GDP.

Conventional wisdom holds that increasing GDP is the key to reducing unacceptably high levels of unemployment. This study has suggested that reducing the debt index may well be the best approach to reducing unemployment levels.

REFERENCES

Bethune, John J., "The Debt Index and its Relation to Economic Activity," *The Journal of Economics and Economic Education Research*, Vol. IV, No 1, 2013, pp. 79-84.

Cebula, Richard, "Budget Deficits, Economic Freedom and Economic Growth in OECD Nations: P2SLS Fixed – Effects Estimates, 2003-2008, *The Journal of Private Enterprise*, Vol. XXVIII, No. 2, Spring, 2013, pp. 75-96.

Duke Website, "Not so Simple Regression Models," http://people.duke.edu/~rnau/simpreg.htm.

EXTERNAL ECONOMIES OF CITY SIZE AND TECHNOLOGY OF PRODUCTION OF MANUFACTURING INDUSTRIES

Farideh A. Farazmand, Lynn University

ABSTRACT

In a highly competitive global market with rapidly changing technology and high demand for new products, transmission of know how, adaptation of new technology, competition for learning innovation and entrepreneurial skills and specialization and outsourcing are of crucial importance for survival of a competitive firm. Economic diversity and specialization of urban areas make the exchange of ideas, the transmission of knowledge and efficient production possible for firms in that area. The principal aim of this paper is to analyze the impact of external scale economies of city size on technology of production of industries.

The study takes a production function approach and examines the impact of external economies of urbanization on the elasticity of substitution. The study covers 19 two digit SIC level industries in 47 SMSAs in the United States. The findings show that the elasticity of substitution is significantly related to the urbanization economies in half of the industries. The analysis of the results reveals that the urbanization economies are a significant factor in affecting the organization and technology of production of industries within the city. The results confirm the relationship between economic diversity and growth of urban centers and technological innovation.

INTRODUCTION

The purpose of this paper is to examine the effects of urbanization economies on production technology of manufacturing industries. Agglomeration economies are external size factors which affect production costs and technology. Urbanization economies are external economies of scale to the firm and industry while localization economies are external economies of scale to the firm.

Most empirical research on agglomeration economies are based on a production function or a relationship derived from a production function. In this study the effects of urbanization economies on a labor demand equation derived from a production function have been measured. Measuring the effects of agglomeration economies directly from a production function is

problematic because of lack of data on capital stock, a labor demand equation resolves this problem because it does not require data on capital stock.

The research investigates the following question: Do external economies of city size affect the production technology of manufacturing industries within the city? It is hypothesized that agglomeration economies significantly affect elasticity of substitution parameter. In this study, urbanization economies have been defined by the extent of industrialization of each city, availability of business services and population of each city.

The extent of industrialization of a city has been measured by total manufacturing employment of each city. The number of business service firms measures the availability of the business services within the city. The least square regression analysis has been applied to the labor demand equation to test the effects of urbanization economies on the elasticity of substitution.

The data covers 19 two digit industries within 47 SMSAs for 1972, 1977 and 1982. The U.S. Census of Manufacturing publishes the industry data every five years, the latest available data is for the 2007 and the next conducted for the year ending December 2012, will be available in December 2013. One reason for use of the historical data for 1972, 1977 and 1982 was due to the lack of availability of data for 2012 at the time this research was conducted, the available data for 2007 was also considered outdated. Therefore the researcher decided to use the earlier years data to give room for structural changes from 1982 to 2012. This study is part of a broader study which will continue to test the model for 2012 when the data comes out in December 2013.

The study uses a production function and labor demand approach to test the impact of the city size on the parameters of a production function, e.g. elasticity of substitution. Although, the historical data of this paper pertains to earlier years, but the theoretical foundation of the study should confirm the validity of the results. However, the model will be also test for 2012 data to compare and contrast the results of 1982 and 2012. For the current study data has been collected from Census of Manufacturers, Census of Industry, Census of Population and State and Metropolitan Area Data Book.

The analysis of findings in this study allowed the following conclusions: In regards to the technology of production, the elasticity of substitution was found to be significantly affected by the urbanization economies variables in half of the industries. The study also reveals that total manufacturing size of each city and business service availability within each city are independently more important means of agglomeration economies than population.

REVIEW OF THE LITERATURE

Spatial agglomerations or clusters have external economies of same sector businesses and employees for firms within that industry that is called localization economies, while urbanization economies are external economies of total economic and social institutions size of a location

which decease production costs of all firms and industries in that location (Harrison, Kelley & Gant, 1996; Hoover, 1971; Isard, 1956; Weber, 1957).

Localization economies, referred by Harrison et al. (1996) as Static Agglomeration, are economies of scale in production resulting in availability of specialized inputs at lower costs (Harrison, et al. 1996).Urbanization or Dynamic Agglomeration Economies refers to spillover of know how and transmission of knowledge in locations with diverse economic activities that not only reduces the cost of general input but also facilitates technological change resulting in higher productivity and lowers average cost (Harrison et al., 1996; Marra, Carlei & Crociata, 2011).

Specialization, diversity, outsourcing, knowledge spillover and competition for learning and innovation lead to conglomeration of industries and firms and development of efficient growth centers which provide a site for small entrepreneurial firms to commercialize their new ideas and products (Cheshire & Malecki, 2003; Marra et al., 2011; Mittelstaedt, Ward & Nowlin, 2006; Rantisi, 2002). Growth of entrepreneurship and technological change lead to an ever increasing specialization, efficiency and growth of cities and regions with divers firms, organizations, knowledge centers (e.g. research universities) and infra-structure (Marra et al., 2011). According to Glaeser (1998) "96 percent of new products innovations occur in metropolitan areas."

Although, some might propose that with substitution of electronic for face to face communication and transactions the agglomeration economies effects of big cities are in decline. However, external economies of know how and technological innovation spillover of big cities could not be substituted by electronic. Moreover, the spatial economies of big metropolitan areas are big enough to out weigh negative factors such as pollution, congestion, crime and social and economic costs (Glaeser, 1998).

The positive relationship between the degree of urbanity and adaptation of new technology has it's roots in the higher degree of competition and faster exchange of ideas in bigger cities (Beeson, 1987; Glaeser, 1998; Harrison et al., 1996; Marra et al., 2011). Therefore, city size also lowers the cost of production by accelerating the rate of technological progress.

Agglomeration economies also change the organization of production of firms and industries by making supply of intermediate products possible through the market. When the firm or industry purchases intermediate products and drop the internal production, the shape of the total average cost of firm or industry changes (Stigler, 1951). This can also be an indication of technological change. In this study the effects of agglomeration economies on the organization and technology of the production will be tested.

Most of the empirical studies of agglomeration economies are based on measuring a production function, or measuring a relationship derived from a production function. Some of these works measure the effects of agglomeration economies on average productivity of labor by estimating a production function which includes the agglomeration economies variable (s)

(Aberg 1973; Greytak & Blackley, 1985; Henderson, 1986; Kawashima, 1975; Moomaw, 1981a, 1981b,1983a, 1983b; Segal, 1978; Shefer, 1973; Tabuchi, 1986).The present study is also based on a production function approach. It is based on a labor demand equation derived from a constant elasticity of substitution (CES) production function.

Carlino (1978) states that the spatial diffusion of innovation begins in the largest urban areas. Beeson (1987) shows that the rate of technical progress across states is affected by agglomeration economies. Harrison et al. (1996) examine the effects of localization and urbanization economies on the adaptation of new technology by manufacturing establishments. They find the economic diversity of the location is a more important factor "for promotion of adopting innovative firm behavior" than localization economy. Marra et al. (2011) examine the growth of 103 Italian provinces and show the relationship between economic diversity of Italian provinces and their economic growth rate.

The empirical work on effects of agglomeration economies on wages show that wages increase with population size. Sveikauskas (1975) analyzes the higher wages paid in large cities by regressing the overall wage rate of workers in manufacturing industries on the population of each city. He shows that money wages increase significantly with city size. Segal (1978) derives the marginal product of labor from a city level production function and show that city size increases the marginal product of labor and consequently the wage rate.

Fuches (1967) finds significant regional differences in the hourly wage rates which are not attributed to the differences in labor compositions. His results show that the ratio of actual to expected hourly earnings increases as the city size increases. Malpezzi, Kiat, and Shilling (2004) examine the relationship between agglomeration economies and the growth of earnings in U.S. metropolitan areas between 1970 to 1999 period. They find "strong evidence" that there is a positive relationship between growth of the metropolitan size, labor productivity and wages. Their results show urbanization economies as a more relevant factor explaining higher wages of bigger cities than localization economies.

Carlino (1985) states that since the 1970s, the manufacturing sector has been growing more rapidly in smaller cities. Carlino (1985) and Moomaw (1983a) use different ways to test the hypothesis that the bigger cities have become less attractive as locations for manufacturing industries. Their results show less favorable position for bigger cities as location for manufacturing industries. However, Black and Henderson (1999) show that "industries are still highly agglomerated" in the bigger cities. They show that there are different types of "manufacturing cities, service centers and market-center cities".

Black and Henderson (1999) suggest that the urban re-concentration in the biggest cities could be the result of transformation of the U.S. economy from manufacturing to financial and high-tec service economy. They examine 15 industries in high-tec and capital goods for 1963-1992 period and show that all industries are agglomerated and the bigger cities have the largest

share of employment of those industries. They report the high-tec industries as the most agglomerated industries.

The current study is concerned with the relationship between urbanization economies and the characteristics of production represented by the elasticity of substitution. Although the sample is the manufacturing industries, but the results could be tested for the service or high-tec sectors in the future.

METHODOLOGY

To measure the effects of agglomeration economies on the production characteristics a general form of CES production function has been chosen. The choice of a general form of CES production function allows for an elasticity of substitution different from one or zero.

To measure the effects of agglomeration economies on elasticity of substitution the parameters of a labor demand equation derived from the production function have been estimated. The labor demand approach allows us to measure the effects of agglomeration economies on the production function parameters without need for capital data. The effects of agglomeration economies on elasticity of substitution have been tested by inclusion of variables measuring urbanization economies. The CES production function in its general form can be written as following:

$$Q = A\,[dL^{-B} + (1-d)K^{-B}]^{-H/B}$$

Where Q is the output, L is labor input, K is capital input, A is an efficiency parameter or Hick's neutral parameter which changes output proportionally for given quantities of input, d($0 \le d \le 1$) is a distribution parameter and B is related to δ the elasticity of substitution parameter, as follows,

$$\delta = 1/(1+B) \qquad -1 \le B \le \infty \text{ since } 0 \le \delta \le \infty$$

and finally, H is the returns to scale parameter.

The following labor demand equation was derived from the CES production function:

$$L = CW^{-\delta}Q^{\beta}$$

Where C is collection of constant terms and $\delta = 1/(1+B)$ is elasticity of substitution.
The agglomeration economies variables have been incorporated into the production function parameter of the elasticity of substitution. The derived demand equation formulation of the production function with agglomeration variables is as follow:

$$L = C\, w^{-(1+a\ln s+b\ln u)/(B+1)} Q^{H(1+a\ln s+b\ln u)+(B-a\ln s-b\ln u)/H(B+1)}$$

Or,

$$L = C\, w^{-\Upsilon} Q^{\beta 1} Q^{\beta 2 \ln Q} Q^{\beta 3 \ln u} \qquad (1)$$

Where,

$$\Upsilon = (1 + a \ln s + b \ln u)/(B+1)$$

$$\beta 1 = \Upsilon + \beta/[H(B +1)]$$

$$\beta 2 = - a/[H(B +1)]$$

$$\beta 3 = - b/[H(B +1)]$$

and,

Elasticity of substitution is,

$$\delta = \Upsilon = (1 + a \ln s + b \ln u)/(B+1)$$

Where s stands for localization economies and u stands for urbanization economies. In (1) elasticity of substitution has been allowed to be affected by agglomeration economies.

Urbanization Economies Variables

Urbanization economies result from the general level of economic activity in an area. The general level of economic activity of a city is a broad concept. We divide it into two main categories: manufacturing sector and service sector. In this way we will be able to study the effects of total manufacturing size and the size of service sector of each city separately on manufacturing industries. In addition to the size of the manufacturing sector and the size of services sector in each city, the population of each city will be also used as a surrogate for city size.

The importance of interrelations between manufacturing industries are in terms of availability and lower price of intermediate inputs, therefore an external spatial size factor reduces cost of production of manufacturing industry (Czamanski & Czamanski, 1976; Carlino 1978). Kelly (1977), Moomaw (1983) and Czamanski and Czamanski (1976) measure the total manufacturing size by total employment of manufacturing sector in an urban area. This study has also used the total employment of manufacturing sector as a measure of the manufacturing size of an urban area.

Another factor changing the price of intermediate products for manufacturing industries is the size of business services in each city. Business services include advertising, computer services, auditing, consulting, telephone answering, janitorial work and provision of temporary office help. The larger the size of business services in a city the more the specialization in the production of these services and the lower their prices. Therefore, the size of business services in a city affects the availability and price of these intermediate products for manufacturing firms

and industries. Thus, the size of business services in each city is an external size factor which affects the costs of manufacturing industries.

The number of business service firms in each city has been used to measure the size of the service sector in each city. Finally, population has been used as a surrogate to capture the effects of any missing urbanization economies variable if there are any.

Empirical Models for Industry

The labor demand in (1) has been estimated at industry level. For an industry, the industry size is an internal factor. Therefore, at industry level the external size factors are the urbanization economies measured by total manufacturing employment, number of business service firms and population of each city. Therefore, the labor demand equation in (2) has been estimated at industry level:

$$\ln T_{ij} = \ln C_1 - Y \ln W_{ij} + \beta_1 \ln Q_{ij} + \beta_2 \ln Q_{ij} \ln M_{ij} + + \beta_3 \ln Q_{ij} \ln BN_j + \beta_4 \ln Q_{ij} \ln POP_j \quad (2)$$

Where, L is ith industry employment in jth city, W is ith industry wage rate in jth city, Q is ith industry output in jth city, M is total manufacturing employment in jth city, BN is number of business service firms in jth city and POP is population of jth city.

Where:

i = 1……..19

j = 1…….47

Ordinary least square regression has been conducted and (2) has been estimated cross section of 47 cities for 19 separate manufacturing industries for three separate years, 1972, 1977 and 1982. The following hypothesis will be tested:

H *It is hypothesized the estimated values of* β_1 *and/or* β_2 *and/ or* β_3 *in (2) are significantly different from zero. If they are, then urbanization economies affect elasticity of substitution.*

Data

Cross sectional analysis over 47 SMSAs for 19 two digit SIC industries have been conducted. Data covers 3 separate years of 1972, 1977 and 1982. The U.S. Census of Manufacturing publishes the industry data every five years, the latest available data is for the 2007 and the next conducted for the year ending December 2012, will be available in December 2013. One reason for use of the historical data for 1972, 1977 and 1982 was due to the lack of availability of data for 2012 at the time of this research, the available data for 2007 was also

considered outdated. Therefore the researcher decided to use the data for earlier years to give room for structural changes in composition of the U.S. cities and industries from 1982 to 2012. But, the fact that the study uses a theoretical foundation and takes a production function and labor demand approach to test the impact of the city size on the parameters of a production function, e.g. elasticity of substitution, should confirm the validity of the results. However, the study will continue and the 2012 data will be also applied and tested for the current model to compare and contrast the results of 1982 and 2012.

The data on labor employment, wage rate, output, number of firms within each industry, industry sales, the total manufacturing employment of each SMSA, the number of business service firms in each SMSA, the population of each SMSA and density of each SMSA have been collected. Data on labor, wages, output and number of firms were at industry level.

All manufacturing industries data were from the U.S. Census of Manufacturers. The total manufacturing employment of each SMSA collected from the U.S. Bureau of the Census, State and Metropolitan Area Data Book. The number of business service firms in each SMSA collected from the U.S. Census of Service Industry. The population data were either from the U.S. Census of Population or the U.S Bureau of the Census, State and metropolitan Area Data Book. The data on density were from the U.S. Bureau of the Census, State and metropolitan Area Data Book or calculated from population data.

Industry labor employment has been measured by all employees of each industry in each SMSA. Industry output has been measured by value added of each industry. The wage rate was computed by dividing total payroll of each industry in each SMSA by all employees of each industry in each SMSA as follows:

W_{ij} = Total payroll of the ith industry in the jth SMSA/All employees of the ith industry in the jth SMSA

RESULTS OF EMPIRICAL INVESTIGATION

Regression analysis was applied to equation (2) to test for non-zero β_2 and/or β_3 and/ or β_4 measuring the effects of city size on the elasticity of substitution parameter of the manufacturing industry. The industry labor demand equation in (2) was estimated for 19 manufacturing industries across 47 SMSAs for three separate years 1972, 1977 and 1982. The number of tables containing the results of estimated values of Υ, β_1, β_2, β_3, β_4 are too long to be listed in this paper. In general, the results were quite good. That is the R^2s were high and the F-test indicated that the specified relations were significant.

As explained earlier, agglomeration economies increase the availability and lower the price of intermediate products, consequently firms and industries drop the internal production of

intermediate products. This changes the organization of the production. It has been also supported by the literature that the rate of technological change is affected by city size. Theoretically, the change in elasticity of substitution of labor and capital is indicative of technological change (Hildebrand & Liu, 1957).

The results of regression analysis of estimating labor demand equation in (2) for 1982 indicated that β_2 , the coefficient of total manufacturing size, was significantly different from zero for five industries, β_3 , the coefficient of business service size was significantly different from zero for two industries and β_4 , coefficient of population in each city, was significantly different from zero for one industry. All together β_s were significantly different from zero for eight industries in 1982.

For 1977, β_2 was significantly different from zero for one industry, β_3 for two industries and β_4 for three industries. Altogether β_s were significantly different from zero for six industries. Results of regression analysis of estimating demand equation in (2) for 1972 showed that β_2 was significantly different from zero for six industries. β_3 was significantly different from zero for one industry and β_4 for three industries. All together 3s were significantly different from zero for eight industries.

The above results indicate that urbanization economies affect the elasticity of substitution of labor and capital for some industries. This could be indicative of the relationship between city size and technological change for some industries. Also, the size of total manufacturing of each city was a more significant variable than the size of business services and population of each city in affecting the elasticity of substitution. The later result could be related to the availability of general intermediate products, out sourcing, transmission of new knowledge, adaptation of new technology and the role of inter industry relationship in each industry's production technology.

The above argument also holds for business services. With city provision of business services internal production is abandoned. Further, the inter industry relationships requires adaptation of each industry to any technological changes arising in another industry. In other words, it could show spread of know how from one industry to another. Population size affected elasticity of substitution for only a few industries. This could also show the effects of bigger cities on technological change.

Table (1) is a summary of results of this paper.

Table 1
SUMMARY OF RESULTS

Year	β_2	β_3	β_4
1982	5	2	1
1977	1	2	3
1972	6	1	3

*Number of β's different from zero. Nonzero β means agglomeration economies affect the elasticity of substitution. β_2 is the coefficient of total manufacturing size, β_3 is the coefficient of business service size and β_4 is coefficient of population.

SUMMARY AND CONCLUSION

Agglomeration economies have been an issue in urban economies and location theory since Van Thunen (1926) and Weber (1957). Discussion of these economies was primarily conceptual until the late 1950s. At that time Vernon (1960) completed the detailed analysis of inter-industry relations and external factor influencing manufacturing in New York States. Later Chinitz (1961) contrasted Pittsburgh's and New York's economies in a way which clearly revealed many of the links between the character and growth of cities and their industrial structures. The next wave of empirical work on agglomeration economies appeared in the 1970s and has continued to this day. The purpose of this research has been to extend this line of research.

The analysis builds on its predecessors in a number of ways. First, production function in which measure of various agglomeration factors were included were estimated for a cross-section of 45 large U.S. cities. Second, the estimates were obtained for two digit SIC manufacturing industries at three different time periods.

The elasticity of substitution was found to be significantly related to one or another of the agglomeration variables in half of the industries in all time periods. Related findings also suggest that there is biased in the use of population as a summary surrogate measure of agglomeration economies. The introduction of the business service availability variable and the incorporation of the localization variable provide a means of considering the use of population as a surrogate variable. In general, the analysis indicates that industry size and business service availability were independently more important than population size.

REFERENCES

Aberg, Y. (1973). Regional productivity differences in Swedish manufacturing. *Regional and Urban Economics.*3(2):131-155.

Black, D. & V. Henderson (1999). Spatial evaluation of population and industry in The United States. *The American Economic Review*. 89.2, 321-327.

Beeson, P. (1989). Total factor productivity growth and agglomeration economies in manufacturing, 1959-73. *Journal of Regional Science.*27(2) 183-199.

Carlino, G. (1978). *Economies of scale in manufacturing location,* Boston: Martinus Xijhoff Social Science Division, Leiden Press.

Carlino, G. (1979). Increasing returns-to-scale in metropolitan manufacturing. *Journal of Regional Science*. 19(3):363-372.

Carlino, G. (1982). Manufacturing agglomeration economies as returns-to-scale: A production function approach. *Papers of the Regional Science Association.* 50(1):95-108.

Carlino, G. (1985). *Declining city productivity and the growth of rural regions: A test of alternative explanation*. Research Department, Federal Reserve Bank of Philadelphia.

Cheshire, P. C. & E. J. Malecki, (2003). Growth, development, and innovation: a look backward and forward. *Papers in Regional Science*. 83.1, 249-267.

Chinitz, B. (1961). Contrasts in agglomeration: New York and Pittsburgh. *American Economic Association*. May: 284-285.

Czamanski, D. Z. & S. Czamanski (1976). Industrial complexes: their typology, structure and relation to regional development. Paper *European Meeting Regional Science Association,* Copenhagen.

Dhrymes, P. (1965). Some extension and tests for the CES class of production function. *The Review of Economics and Statistics* .47(4):357-366.

Ferguson, C.E. & T.R. Saving (1969). Long-run scale adjustment of a perfectly competitive firm and industry. *The American Economic Review*.59(5):774-783.

Fogarty, M.S. & G.Z. Garofalo (1978). An exploration of the real productivity effects of cities. *The Review of Regional Studies, Spring.* 8:65-82.

Fogarty, M.S. & G.Z. Garofalo (1980). Urban size and the amenity structure of cities. *Journal of Urban Economies.* 8(3):350-361.

Gershung, J.I. & I.D. Mills. (1983). *The new service economy*. New York: Praeger Publishers.

Glaeser, E. L. (1998). Are cities dying? *The Journal of Economic Perspectives.* Spring, 12(2):139-160.

Goldstein, G.S. & T.J. Gronberg, (1984). Economies of scope and economies of agglomeration. *Journal of Urban Economics.* 16(1):91-104.

Greytak, D. & P. Blackley (1983). A firm analysis of returns-to-scale in metropolitan manufacturing: implications of agglomeration economies. Presented to the *Thirtieth North American Meeting of the Regional Science Association.*

Greytak, D. & P. Blackly (1985). Labor productivity and local industry size: further issues in assessing agglomeration economies. *Southern Economic Journal.* 51(4):11-21.

Harrison, B., M. R. Kelley & J. Grant (1996). Exploring the intersection of agglomeration, firm effects, and technological change. *Economic Geography*, 72.(3):233.

Henderson, V. (1986). Efficiency of resource usage and city Size. *Journal of Urban Economies*. 19: 47-70.

Hoch, I. (1972). Income and city size. *Urban Studies*. 9: 299-328.

Hoover, E. (1971). *An Introduction to regional economies*. New York: Alfred Knoff.

Isard, W. (1956). *Location and space-economy*. New York: John Wiley and Sons, Inc. and The Technology Press of Massachusetts Institute of Technology.

Isard, W. (1975). *Introduction to regional science*. Englewood Cliffs, New Jersey: John Wiley and Sons, Inc. and The Technology Press of Massachusetts Institute of Technology.

Kawashima, T. (1975). Urban agglomeration economies in manufacturing industries. *Papers of the Regional Science Association.* 34(1):157-172.

Kelley, K. (1977). Urban dis-amenities and the measure of economic welfare. *Journal of Urban Economics*. 4(4):379-388.

Leven, C. (1978). *The mature metropolis*. New York: Lexington Books.

Malpezzi, S., S. Kiat-Ying & J. D. Shilling, (2004). It is what we do on how we do it? New evidence on agglomeration, economics and metropolitan growth. *Real Estate Economics*. 32(2):265-295.

Marcus, M. (1965). Agglomeration economies: A suggested approach. *Land Economics*. 41(3):279-284.

Marra, A., V. Carlei, & A. Crociata, (2011). Economic diversity and employment levels in Italian provinces over the period 1951-2001; Does a recurrent industrial pattern exist? *Modern Economy*. 213(July):438-445.

Meras, K. (1973). On the urban agglomeration and economic efficiency. *Economic Development and Cultural Change.* 21(2):309-324.

Meyer, D. (1977). Agglomeration economies and urban industrial growth: A clarification and review of concept. *Regional scienceperspectives.*7(1):80-92.

Mittelstaedt, J. D., W. A. Ward & E. Nowlin (2006). Location, industrial concentration and the propensity of small US firms to export: Entrepreneurship in the international marketplace. *International Marketing Review*. 23(5):486-503.

Moomaw, R.L. (1981a). Productive efficiency and region. *Southern Economic Journal.* 48(2):344-357.
Moomaw, R.L. (1981b). Productivity and city size: A critique of the evidence." *Quarterly Journal of Economics*. 96(4):675-688.
Moomaw, R.L. (1983a). Spatial productivity variations in manufacturing: A critical survey of cross-sectional analysis. *International Regional Science Review.* 8(1):1-22.
Moomaw, R.L. (1983b). Is population a worthless surrogate for business agglomeration economies? *Regional Science and Urban Economics.* 13: 525-545.
Morony, J.R. (1972). *The structure of production in American manufacturing*. Chapel Hill, NC: The University of North Carolina Press.

Rantisi, N. M. (2002). The competitive foundations of localized learning and innovation: The case of Women's garment production in New York City. *Economic Geography*. 78.4: 441-462 (Oct)
Schmenner, R. W. (1978). The manufacturing location decision: Evidence from Cincinnati and New England. *Report of the Office of Economic Research, Economic Development Administration, U.S. Department of Commerce*.
Segal, D. (1978). Are there returns-to-scale in city size? *The Review of Economics and Statistics*. 58(3):339-350.
Shefer, D. (1973). Localization economies in SMSA's: A production function analysis." *Journal of Regional Science.* 13(1):55-64.
Stanback, T.M. (1979). *Understanding the service economy*. Baltimore: The Johns Hopkins University Press.
Stigler, G. (1951). The Division of Labor is Limited by the Extent of the Market. *The Journal of Political Economy*.
Sveikauskas, Leo. (1975). The productivity of cities. *Quarterly Journal of Economics*. 89(3):393-413.
Rodgers, J. 1976. Theory of the firm: A simplification and synthesis of recent contributions. *The Eastern Economic Association.* 3(2):211-228.
Tabuchi, T. (1986). Urban agglomeration, capital augmenting technology, and labor market equilibrium. *Journal of Urban Economics.* 20(2):211-228.
Census of Manufacturers. (1972). U.S. Bureau of the Census.
Census of Manufacturers. (1977). U.S. Bureau of the Census.
Census of Manufacturers. (1982). U.S. Bureau of the Census.
Census of Service Industry. (19727). U.S. Bureau of the Census.
Census of Service Industry. (1977). U.S. Bureau of the Census.
Census of Service Industry. (1982). U.S. Bureau of the Census.
Census of Population. (1972-1982). U.S. Bureau of the Census.
Von Thunen, J. H. (1826). *Per Isolierte Staat in Beziehung auf Land irtschaft und Nationalokcncmie* , English translation Hall, P. (ed.) 1966, *Von Thunen's Isolated State* , Oxford: Pergamon.
Vernon, R. (1960). *Metropolis 1985*. Cambridge, Mass.: Harvard University Press.
Weber, A. (1957). *Theory of the location of industries*. Chicago: The University of Chicago Press.

ALIGNING ECONOMICS PROGRAMS WITH AACSB ACCREDITATION PROCESSES

Laura E. Fitzpatrick, Rockhurst University
Cheryl McConnell, Rockhurst University

ABSTRACT

A challenge facing business schools, and of particular interest here, economics programs in business schools, is that of aligning programs to be consistent with the assessment expectations for AACSB accreditation. In the process of defining expectations and measuring achievements, a torrent of new vocabulary, processes, and expectations on faculty have been imposed. Many faculty members feel overwhelmed and resentful about the process and requirements. However, what can await a school and/or program at the end of the process is a unified, articulable view of program learning goals, how the program seeks to achieve the goals, and whether the goals are being met. The authors present a systematic process by which an economics program was successfully aligned with AACSB processes and standards, and examples of assessment plans, reports, and outcomes are provided. Excellence in student learning is the goal the authors share with others in their profession, and aligning programs as described here can create an opportunity to determine where quality learning is already happening, and where changes may be needed in order to achieve this level of excellence. It is hopeful that by describing the requirements and implementation processes of an assessment plan, this can serve as a model for others who engage in the process.

INTRODUCTION

Beginning and sustaining program level assessment can be a daunting task. Economics programs in AACSB accredited business schools often are either required or requested to have robust assessment programs in place to demonstrate continuous improvement processes. In many ways, these expectations take schools in a very useful, but simultaneously, overwhelming direction. They drill down to what a program is meant to be and if that program is achieving what it seeks to achieve. In the process of defining expectations and measuring achievements, a torrent of new vocabulary, processes, and expectations on faculty have been imposed. The key to success is to clearly understand expectations, and to develop achievable assessment and reporting processes. The purpose of this paper is to identify expectations and provide actual examples of successful accreditation processes at a recently reaccredited AACSB business school.

AACSB standards clearly define the broad steps required of an assessment program, and state that the assessment process must be faculty driven. According to the AACSB Assessment Resource Center,

> *"The standards call for schools to define learning goals, assess student achievement for these goals, and utilize what is learned through assessment to continually improve their curricular programs."*
>
> *"Faculty involvement in, and ownership of, the assurance of learning process is critical. Faculty are expected to be actively involved in all stages of the assessment process including defining goals, curriculum alignment, developing appropriate measures, implementing course-embedded measures, and, improving the school's curriculum." (AACSB Assessment Resource Center, 2011)*

The implications of these standards are that there must be clarity in what the school seeks to do, how it continually seeks excellence, and that this cannot be a delegated job to one or two individuals or faculty members. The economics program assessment processes described in this paper provides concrete examples and documents from which other programs can use to build their own mission-based assessment process.

DEFINING LEARNING GOALS FOR THE PROGRAM

AACSB expectations with respect to learning goals are that:

1. Learning goals should link to the mission; thus, learning goals will differ from school to school.
2. Learning goals translate the more general statement of the school's mission into the specific educational accomplishments expected of its graduates.
3. Learning goals must be defined for each program. Departmental goals and/or course goals (which are not required by AACSB) are not a substitute for program goals.
4. Learning goals must include both general and management-specific knowledge and skills.
5. Four to ten goals should be developed for each program. Schools are not required (or even encouraged) to develop and assess learning goals for all of the knowledge and skills areas listed in [AACSB] Standards 15–21. (AACSB Assessment Resource Center, 2011)

Therefore, the first step in the assessment process is the establishment/definition of the learning goals for the program. By requirement, the learning goals for any program need to reflect the mission of the school. This can be a lengthy process – not because a program is unaware of what it seeks to be or because articulating that is hard, but rather the challenge can come in the need for levels of concurrence of learning goals across the institution. For example,

the BA in Economics program—which is the model here—is one of two undergraduate programs (the other being a BSBA) in the Helzberg School of Management at Rockhurst University. The university has a mission, the Helzberg School of Management has a mission and learning goals in the BA in Economics program need to be consistent with all of them.

The Helzberg School of Management's approach to this process was to begin with the establishment of general learning goals with an eye to one program (BSBA), gain faculty approval, and then extend the process to other programs in the school (BA in Economics, MBA, and Executive MBA). A small group of faculty members representing the various disciplines in the school met to define learning goals for the undergraduate business school BSBA degree. An eye on this one program helped move the process forward as it was less abstract than trying to establish the goals for all programs at once. Drafts were shared with the faculty at large. Input was received on the goals themselves and whether they captured what needed to be generalized for the entire school. After revisions and a faculty vote, six learning goals were decided on for the BSBA. These were then generalized for the entire Helzberg School of Management. The six learning goals of the Helzberg School of Management fall under the themes of: Leadership; Ethical Behavior and Corporate Social Responsibility; Business Skills and Knowledge; International/Global; Information Analysis and Application; and Communication. These fundamental learning goals became the starting point for other programs—including economics—to establish theirs while retaining concurrence within the entire university.

Each program then articulated their corresponding program level learning goal that reflected the nuances, depth, and focus of these goals for each particular program. This gave a unified focus through the school while developing the particular profile of each program.

The Helzberg School of Management was at the forefront of this process in the larger university. Their learning goals were created with the mission statements of the Helzberg School and Rockhurst University in view, but the goals also then served as a starting point for when, several years later, the entire university began the process of establishing university level learning goals. The result was concurrence throughout the institution. Table 1 shows this concurrence from the university through the business school to the economics program.

Although these goals have been established, they remain dynamic documents. Reexamination of the Helzberg School of Management learning goals is undertaken both systematically and in an ad hoc manner as questions or needs are presented. A recent example of this was found in the executive MBA program where revision of some learning goals was made to better reflect the desired outcomes of the program. After approval by the appropriate program committee, the revisions were presented and voted upon by the entire faculty. As such, the learning goals retain a vibrancy and progression that is so necessary for continuous improvement, and it allows the business school to respond as necessary to changes in environmental and strategic factors.

Table 1: LEARNING GOAL CONCURRENCE ACROSS UNIVERSITY AND ECONOMICS PROGRAM		
Rockhurst University	**Helzberg School of Management**	**BA, Economics**
Leadership *The commitment to develop the gifts and talents of self and others to make a positive difference in the world*	Leadership *Demonstrate the pursuit of personal excellence while helping others develop to their full potential*	Leadership and Public Policy • *Demonstrate leadership skills through formulation and evaluation of beneficial public policy* • *Demonstrate leadership skills through educating others about public policy*
Ethics and social justice *The commitment to create a more just world and to live with integrity, humility, tolerance, and empathy*	Ethical Behavior and Corporate Social Responsibility *Analyze ethical and corporate social responsibility issues in context and implement appropriate action(s)*	Ethics and Social Justice • *Distinguish and apply both positive and normative economic tools to define and debate economic issues and policy.* • *Recognize and analyze issues relating to personal ethics and social justice to propose and defend courses of action to create a more just world.*
Academic knowledge *The capacity to assimilate and apply a broad range of skills, knowledge, and abilities to a chosen field of study*	Business Skills and Knowledge *Explain, integrate and apply foundational business knowledge and skills to effectively lead and manage organizations*	Economics Skills and Knowledge • *Define, describe, demonstrate, and apply intermediate level economic theory.* • *Apply scientific method to develop new knowledge*
International and cultural understanding *The appreciation of cultural differences and commonalities, and the ability to interact with sensitivity and alertness as citizens of the world*	International/Global *Demonstrate the achievement of a global perspective that encourages participation in the complex, integrated world-wide business community*	International/Global *Integrate relevant cultural, social, political, historical, geographic, and environmental factors into the analysis and debate of economic issues and courses of action.*
Critical and creative thinking *The ability to search for knowledge, investigate questions, and apply information systems in a discerning and innovative manner*	Information Analysis and Application *Identify, access, analyze and synthesize appropriate business information*	Critical Thinking and Information analysis/application *Identify, access, and analyze relevant quantitative and qualitative information to evaluate economic issues/problems, to develop forecasts, and to select and evaluate appropriate courses of action*
Communication *The ability to communicate effectively in a variety of contexts and with awareness of purpose and audience*	Communication *Communicate effectively, and create an environment where effective communication can occur*	Communication *Produce and deliver effective written products and oral presentations in a variety of contexts using effective technologies*
Self formation *The discovery and cultivation of spiritual, physical, social and emotional well-being*	It was an explicit decision that this learning goal is primarily achieved and assessed through the extra-curricular areas of the university.	

ASSESSING STUDENT ACHIEVEMENT OF LEARNING GOALS

After program level learning goals are established, the next step is to assess whether students have achieved the learning goals by the end of their program. AACSB's specific expectations for assessment of student learning are that:

1. Student performance on learning goals must be assessed systematically and routinely. No one approach to assurance of learning is prescribed. Assessment programs should include direct measures of learning. Course grades are not program assessment measures.
2. Program assessment does not require that every student be assessed. Sampling is acceptable as long as an appropriate and representative sampling methodology is utilized. (AACSB Assessment Resource Center, 2011)

The authors find that the assessment of student learning step, more than the others, troubles and intimidates faculty members. DeMoranville notes three broad reasons faculty members resist assessment requirements. First, faculty members are too busy with current responsibilities in teaching, service, and scholarship and therefore have little time for activities they view as busy work. Second, they question the true value to be gained through assessment with the high costs of additional work accruing to the faculty and the potential benefits of better learning accruing to the students. Lastly, they are concerned about potential limitations on their ability to design and deliver courses as they desire. (DeMoranville, 2010, pp.24-25) Perhaps this is why Lederman noted that a 2009 survey by the National Institute for Learning Outcomes Assessment found that "campus leaders considered involving faculty in assessment to be one of their greatest challenges." (Lederman, 2011) Responses the authors have received to the need to plan and perform assessment range from an unwillingness to learn a new system and take on new responsibilities, to nervousness of being placed under the microscope in teaching. Concerns must be understood, and an assessment process designed that is manageable, can provide meaningful information about achievement of student learning, and will continue to allow faculty members to design and deliver their courses in ways they believe are appropriate.

Creating the Assessment Plan

Creating the assessment plan begins with setting a timetable for assessing program level learning goals. A multi-year plan that rotates the assessment of one or two goals per year makes it very achievable. For Economics at Rockhurst University, the plan was created by looking at the entire curriculum, selecting courses (based on existing course learning objectives) in which it made sense to assess the program learning goals, and spreading these assessments out over time. The plan avoids assessment overload in one particular course, and it allows for the establishment of baseline and end-of-program assessment in a systematic way. For example, according to the plan, data for ethical analysis is collected from the Developing World course, analyzed and reported to the economics faculty members in year one. Any recommendations for changes made in courses tied to assessment are approved in that academic year. In year two of that learning goal's assessment cycle, any recommended changes are implemented in the course(s) where changes were recommended and approved. In year 3, data collection, analysis, and recommendations will happen again to determine whether standards were met after changes were

implemented. The staggering of different goals to be assessed in different years greatly simplifies the assessment process. Note that this approach is entirely consistent with AACSB expectations:

> *"AACSB standards specify "a systematic process" only. Each goal does not have to be assessed every year, but a systematic process is needed to insure all goals are assessed to support meaningful curricular change and development. Normally, each goal should be evaluated at least twice over a five year AACSB review cycle." (AACSB, 2007, p 15)*

Assessment Methodologies

Once the rotation plan has been established, an appropriate assessment methodology for each learning goal needs to be determined. Although methodologies may include indirect techniques such as surveys, interviews and focus groups, they must include direct measures of student learning such as assignment artifacts or assessment exams. Assessment at Rockhurst University is based primarily on the direct assessment methodology of course embedded assessment. This is a university wide emphasis, so it clearly fits into the organizational culture and satisfies accreditation expectations. Course embedded assessment uses existing course assignments, activities, papers, and/or exam elements to directly assess student learning.

According to McConnell et al., a well-designed course embedded assessment methodology identifies student artifact that provide evidence of the learning goal, and matches it with an appropriate measurement technique that allows faculty members to determine whether a learning standard has been achieved. Table 2 below describes common course-embedded artifacts and the related measurement techniques that a faculty member might use. (McConnell, Hoover, and Miller, 2008)

Table 2: Artifacts and Measurement

Course-Embedded Artifacts	Measurement Techniques
Multiple choice exam questions related to a particular learning goal	Percent correct, analysis of incorrect responses
Short-answer exam questions, essays, research papers	Level of achievement rubrics
Oral presentations	Oral presentation rubric
Case study reports	Case study rubric
Lab performance	Skills checklist

In the economics program, all of these techniques to capture student learning—with the exception of skills checklists for lab performance—have been used. For ease of use and consistency in data comparison, standardized rubrics for a number of the program learning goals were created, tested, and adopted. Rubrics are useful any time students are making a non-objective response because they clarify the dimensions to be graded and provide scales or descriptors of student performance (McConnell et al., 2008). The common rubrics used by the economics faculty are for the learning goals of communication (both oral and written), ethics, global, critical thinking, and information analysis/ application. An example of the standardized

critical thinking and information analysis/application rubric for Economics is included in Table 3 below for reference.

Table 3: Critical Thinking and information Analysis / Application Rubric

Learning Objective	1 - Exceptional	2 - Superior	3 - Commendable	4 - Rudimentary	5 - Minimal
Identify	Demonstrates a clear/accurate and comprehensive understanding of data collection process and statistical theories and concepts.	Demonstrates an adequate understanding of data collection process and statistical theories and concepts.	Demonstrates a general understanding of data collection process and statistical theories and concepts.	Demonstrates an incomplete understanding of either the data collection process or statistical theories and concepts.	Demonstrates incomplete or mostly incorrect understanding of data collection process and statistical theories and concepts.
Access	Presents a concise and correct explanation for choosing particular techniques and models to fit and forecast the data.	Presents a correct explanation for choosing particular techniques and models to fit and forecast data.	Presents an acceptable explanation for choosing particular techniques and models to fit and forecast data.	Presents a limited and partially incorrect explanation for choosing particular techniques and models to fit and forecast data.	Presents no or completely incorrect explanation for choosing particular techniques and models to fit and forecast data.
Analyze	Excellent use of statistical evidence and prior knowledge (of topic) to compare models' performance and to make recommendations for future forecasts.	Comparisons and recommendations are based on appropriate and correct statistical evidence and prior knowledge.	Comparisons and recommendations are based on mostly appropriate or correct statistical evidence and prior knowledge.	Comparisons and recommendations incomplete and/or selection of preferred model are based on inappropriate or incorrect statistical evidence and prior knowledge.	No attempt to compare models' performance and/or to make recommendations for future forecasts.

The advantage of developing standardized rubrics for program learning goals is found in the process of its creation and application. The creation of a rubric involves faculty collectively discussing and determining which dimensions and scales are important for their program, and expressing them in a concise and communicable way. The result is cohesion among faculty on student achievement expectations. Some might voice concern that there may be elements that one individually faculty member values highly that does not make the final rubric, but in practice, this is not a problem. For assessment purposes, an instructor who is gathering data for an assessment must use the common rubric dimensions *at a minimum*. Additional dimensions geared toward a particularly desired outcome(s) in a course or an assignment can easily be added to the rubric, but only those designated for assessment need be part of the formal data analysis and recommendation process. This flexibility preserves great freedom for the faculty member

while providing essential assessment information to determine whether program learning goals are being met.

In the data collection and analysis phase, rubric use is also very helpful because the standardized rubric makes the essential connection between learning goals and assessment results (Ammons and Mills, 2005). A standardized rubric can be used for student assignments in multiple courses, and it allows comparisons between baseline course assignments and end-of-program assignments. For example, the information analysis / application rubric shown in Table 3 can be used in a sophomore level statistics course to determine a baseline level for incoming student performance, and then again in a capstone course for an end-of-program assessment of student learning.

USING ASSESSMENT RESULTS

The final step in program assessment involves feeding conclusions and recommendations that flow from the assessment data back into the program for continuous improvement. The purpose of assessment is not the gathering of data or the creation of more work for the faculty member; it is to identify an act on areas that need improvement or attention. This view is emphasized by Banta who states,

Outcomes assessment is simply not worth doing unless it is used to enhance the students learning experience—by improving instruction in a single class, the structure or sequencing of a curriculum, or the process of offering student services that complement coursework." (Banta, 2005, p. 38)

AACSB has the specific expectations that:

1. Assessment results must be analyzed, disseminated, and utilized by the faculty toward curriculum planning.
2. For initial accreditation and accreditation maintenance purposes, schools will be expected to define their learning goals conceptually and operationally, discuss how they are addressed in the curriculum, and demonstrate levels of student achievement for each goal. Schools also will be expected to show how assessment results subsequently impacted their curriculum planning. (AACSB Assessment Resource Center, 2011)

This step of applying changes to the curriculum for improvement is what AACSB calls 'closing the loop' on a round of the assessment process. In the authors' last AACSB site visit, the visitation team stressed the expectation that institutions not only assess, but make recommendations and act on the recommendations. The loop is not considered to be closed in a learning goal assessment until any recommended changes are implemented. As the cycles are repeated over time, the faculty can not only determine whether changes implemented produced the desired outcome, but also if there are additional areas in need of improvement.

Assessment Reports

Documenting and tracking the assessment plan requires simple reporting and archiving. In the economics program, assessment reports are created whenever a program learning goal is assessed. The report summarizes the relevant data collection information (learning goal, where assessment took place, results, and recommendations), includes the faculty member's recommendations for improvement, and indicates where the electronic copy of the report and archived artifacts of student learning can be found. The Helzberg School of Management has found that the most efficient way to maintain these required archives is in electronic form on a dedicated drive.

Assessment reports are presented at economics departmental meetings where all faculty members determine what the final recommendations will be and what, if any, changes are to be made to the assessed course or other courses in the program. An example of such a report is found in Table 4 below. The following year, the changes are applied and the courses await the next round of assessment.

Annual Reports and Cumulative Assessment Plan Reports

Although the loop of assessment is closed for a learning goal once assessment results are discussed and any changes are implemented, documentation of the assessment plan is essential, especially for AACSB and regional accreditation bodies.

In the last AACSB site visit rotation, the Helzberg School of Management instituted a summary annual reporting process to assure that it captured all assessment activities in each program, and to also keep track of the accumulated assessment activities per learning goal. This way, major comprehensive assessment reports do not need to be created for site visits or self-studies, but rather the current state of assessment in every program is updated and documented each year.

Program chairs prepare an annual assessment report summarizing all of the program assessments, discussions, changes, and pending plans. An example of the annual assessment report for Economics can be found in Table 5. As Table 5 shows, the annual report details the learning goals assessed in the year, the observations and discussions, and the recommendations and interventions. The distinction between the Assessment Report in Table 4 and the Annual Report in Table 5 is that Assessment Reports are prepared by individual faculty members performing course-embedded assessment for a single learning goal, and the Annual Report is prepared by the program chair, and it documents the departmental discussions and decisions about all learning goals assessed in the academic year.

The final step is to add to a Cumulative Assessment Plan Report (standardized across all programs in the Helzberg School of Management) that accumulates all of the varied assessment loops for a program over time. A sample of the table for the BA in Economics is found in Table

6. Just looking at the table can be daunting, but once created, all that need be done to the table each year is adding the few new lines of assessment information under learning goals that have been assessed. When it comes time to contribute assessment information to accrediting bodies such as AACSB or a regional accrediting body, the documents are up-to-date and assessment of learning goals and cycles can easily be viewed and shared.

Table 4: Sample Learning Goal Assessment Report

ASSESSMENT REPORT
Ethics and Social Justice
Laura Fitzpatrick

GOAL	DATA	ANALYSIS	OBSERVATIONS/RECOMMENDATIONS
Ethics and social justice Recognize and analyze issues relating to personal ethics and social justice to propose and defend courses of action to create a more just world.	EC3400-DEVELOPING WORLD Econ majors, core SRII students (business overlap), Global Studies, junior/senior 5 page case concentrating on economic policy, ethics, and CSR Potentially first case in ethical analysis	1. Students had most difficulty with the performance dimension that required stakeholder analysis and implication of courses of action. 2. Students had least difficulty with the performance dimension that required the recommendation and support of a course of action 3. Student performance percentages are available on attached rubric templates	1. Rubric worked well for exercise. 2. Achievement standards should target 90% acceptable performance or above on each of the four performance dimensions. 3. Document more explicit ethics related learning objectives. Course did not include a learning objective solely tied to ethics although it was a measurable component of the course and was even inferred in existing learning objectives. 4. Explicitly introduce stakeholder analysis exercises

Note: In this space, one would indicate where the supporting data analysis and the archives of student work can be found.

Table 5 : Annual Assessment Progress Report

Program: BA, Economics
Program Coordinator: Prof. Laura Fitzpatrick
Inclusive dates:

Overview
Assessment this year included data collection, analysis, and recommendations in global/international and rubric development, piloting, and recommendations in critical thinking and information analysis/application. Scheduled rubric development or modification in economic skills and knowledge and in leadership and public policy has not yet been completed. There is additional global/international assessment that has yet to be completed as well.

Data Analysis: Global/International
The preliminary review of the existing **'**global/international' assessment data was discussed in the BA in Economics Curriculum and Assessment Committee (BACAC) meeting on August 17, 2011 and the initial recommendation was adopted. Highlights of those recommendations are listed below while the full report and corresponding artifacts can be found on the HSASSESS drive, program assessment, BA, EC3400 F10 L Fitzpatrick.

The global analysis rubric segment was reviewed and approved by the BACAC committee for use in Undergraduate program assessment. The BACAC determined that 85 % of students scoring acceptable or above would be the targeted achievement level for each performance dimension of the above rubric.

Observations & analysis:
Students did meet the targeted competency level in all dimensions of the rubric. Actual achievement was 100%. Students are achieving targets in this area. The area of weakness in the course was not found in this goal but, rather, in the ability to apply different theoretical perspectives. This is currently an assessment focus for SR courses through the modal group.

Recommendations:
At this point, we are looking to gather more data to see if there is consistency across courses and we are initiating assessment of this goal at the introductory level. This should give us a better view of the goal throughout the program as well as potential areas of intervention.

Data Analysis: Critical Thinking and Information Analysis/Application
The preliminary review of the existing **'**Critical Thinking and Information Analysis/Application' assessment data was discussed in the BACAC meeting on August 17, 2011 and the initial recommendations were adopted. Highlights of those recommendations are listed below while the full report and corresponding artifacts can be found on the HSASSESS drive, program assessment, BA, EC4001 S11 X. Pham.

The critical thinking and information analysis/application rubric was adapted from that of the BSBA program to customize it to the BA program. This revised rubric was reviewed and approved by the BACAC committee for use in Undergraduate program assessment.

The BACAC determined that 85 % of students scoring acceptable or above would be the targeted achievement level for each performance dimension of the above rubric.

Observations & analysis:
Students did meet the targeted competency level in all dimensions of the rubric. Actual achievement was 92.3%.
Students performed better on the new topics introduced in the course than on those that required retention from topics in BUS2200. This should be addressed.
Students were relatively weaker on background knowledge of data sets they chose.
Technical problems (frequent crashes) with the Excel forecasting add on created a great deal of frustration and unnecessary challenges not related to subject matter for students.

Recommendations:
Some kind of retention focused efforts from BUS2200 would benefit students. Instructor has indicated a desire to focus on course interventions to increase understanding of background knowledge of data sets. Strong recommendation of alternative forecasting tool that will not be a distraction to learning. Something such as SPSS would also be marketable from a student perspective.

Interventions:

1. BACAC is currently brainstorming what approach would be best to help retention desired. Some options are use of a primer students retain, use of Livetext to create a resource bank of BUS2200 material that can be revisited, and additional review. The final decision will be made and made Fall 2011 for implementation in the next offering of the course, Spring 2012.
2. BACAC is investigating the cost and feasibility of an alternative forecasting tool to be used in the course. The final decision will be made and made Fall 2011 for implementation in the next offering of the course, Spring 2012.

Table 6: Sample of BA Summary Assessment Results Table

BA SUMMARY ASSESSMENT RESULTS

Learning Goal	Academic Year	Assessment and Results	Loops Closed
Ethical Behavior and Corporate Social Responsibility	AY 08/09	• Rubric developed and piloted in AC4750 and EC3400. Results indicated rubric needed enhanced descriptors on multiple dimensions.	Rubric modified and adopted by faculty. 1st loop closed.
	AY 09/10	• Assessment data collected in EC4940 and EC4200. In EC4940 students met 90% standards that had been set in two of four dimensions, and 70% and 80% in remaining dimensions. In EC4200 students met 90% in two and 85% and 80% in two others, but they were not the same low and high scoring dimensions across the courses. With varied results, the recommendation speaks to perhaps instructor specific adaptations in class to increase the achievement levels in weaker dimensions. Faculty started questioning whether 90% is the ideal we seek and whether it sets the proper level of achievement for acceptable performance for program assessment purposes. The recommendation is to examine this and potentially revise target achievements.	2nd loop closed with instructor changes in individual courses. Target competencies modified to 85% satisfactory or better.
	AY 10/11	• Change implementation year.	
Summary: Two loops closed. Third loop begins AY 11/12			

CONCLUSION

An integrated approach to program level learning assessment is no longer a choice that schools face. For myriad reasons, not the least being expectations from accrediting bodies, schools must engage in the process and develop systems that will work for them. Although universities at the early stages perceive the process to be overwhelming, a program need begin with only small steps. Ewell notes that,

> *"The prospect of starting an integrated program of learning assessment can seem overwhelming … but that shouldn't be an obstacle to getting started. Institutions that have built comprehensive, highly integrated, sell-documented systems of assessment have been developing their practices for years. They started with small steps, perhaps with only one course, and worked their way up to the whole." (Ewell, 2003, p.33)*

A framework and models of these beginning steps can be drawn from this paper. The essential conditions for success are that faculty concerns be understood, the assessment process and design is simple and achievable, the process provides meaningful information about student learning, and that most, if not all, faculty members are active participants in the assessment process and discussions.

REFERENCES

AACSB Assessment Resource Center, Overview and frequently asked questions regarding AACSB expectations, Retrieved September 7, 2011, from http://www.aacsb.edu/resources/assessment/faq.asp

AACSB International Accreditation Coordinating Committee, (2007). AACSB Assurance of Learning Standards: An Interpretation*, AACSB White Paper*, 20 November 2007.

Ammons, J. and Mills, S. (2005). Course-embedded assessments for evaluating cross-functional integration and improving the teaching-learning process. *Issues in Accounting Education,* 20(1) February, 1-20.

Banta, T., (2005). How much have we learned? *BizEd., September/October*. Retrieved from http://www.aacsb.edu/publications/archives/sepoct05/p34-39.pdf

DeMoranville, C. (2010). Making sense of assessment. *BizEd*, *March/April*. Retrieved from http://www.aacsb.edu/publications/Archives/MarApr10/22-39F-Assessment.pdf

Ewell, P., (2003). The learning curve. *BizEd., July/August.* Retrieved from http://www.aacsb.edu/publications/Archives/JulyAug03/p28-33.pdf

Lederman, D. (2010). Faculty role in assessment. *Inside Higher Ed., May 29, 2010*. Retrieved from http://www.insidehighered.com/layout/set/print/news/2010/05/28/assess

McConnell, C., Hoover, G., and Miller, G., (2008) Course embedded assessment and assurance of learning: examples in business disciplines. *Academy of Educational Leadership Journal*, 12 (3) 19-34.

POLARIZATION ON ECONOMIC ISSUES OVER TIME – A SURVEY OF DELEGATES TO THE NATIONAL CONVENTIONS

Doris Geide-Stevenson, Weber State University
Nazneen Ahmad, Weber State University
Dan A. Fuller, Weber State University

ABSTRACT

We ask whether partisan polarization on economic issues has increased over time among political elites. Based on survey results of party delegates to the national conventions of the Democratic and Republican parties in 1992, 2000 and 2008, we construct various measures of consensus. The surveys ask delegates whether they agree, agree with proviso or disagree with a number of economic propositions. For propositions common in all three time periods, we compare the level of consensus within and between the two political parties. Our results suggest a divergence of opinion between Republican and Democratic delegations from 2000 to 2009. This divergence of opinion is due to an increase in the level of consensus among Republicans from 2000 to 2009 but mitigated by a decrease in the level of consensus among Democrats from 1992 to 2000. While we confirm diverging opinions between 2000 and 2009, we also find that the 2009 survey results mirror some of the results from 1992, suggesting that the current polarization is not historically unique with respect to economic issues.

INTRODUCTION

Media accounts of the current political climate in the United States often focus on the high degree of polarization between Republicans and Democrats. Such accounts describe a trend of "relentless" and "vitriolic" polarization (Martin, 2010; Economist, Feb. 2010) along with the "death of moderates" in American government (Beinart, 2010). In trying to explain the apparent polarization, wedge issues or views founded in religious and moral values have been found to be important (Glaeser, Ponzetto and Shapiro, 2005, Layman, 1999). However, Glaeser, Ponzetto and Shapiro (2005) note that party platforms on economic issues, as opposed to religious or cultural issues, are less polarized citing language that is "quite moderate and similar across platforms".

This paper explores polarization on economic issues based on a set of economic propositions distributed to Democratic and Republican delegates to the national conventions preceding presidential elections in the years 1992, 2000, and 2008. The benefit of being able to compare survey results from three different time periods puts the current discussions of polarization in a larger context.

Delegate surveys have been used regularly to gather information on the political perspectives of party elites but infrequently focus on economic issues (Miller and Jennings, 1992). This paper amends the existing discussions of polarization by focusing on economic issues. Defining polarization as a divergence of opinion between the two parties, we construct two measures that provide the substance for our study. The first measure, the relative entropy index, is used to indicate the degree of consensus or the convergence of opinion in each party. The second measure, a conditional measure of broad agreement, is used to measure the direction of as well as an indicator of the level of consensus in each party. Results indicate that the average level of consensus among 2009 Republicans is significantly higher than in 2000. Conversely, the average consensus of opinion among 2000 and 2009 Democrats is significantly lower than in 1992. We also find that while opinions of Democrats and Republicans are somewhat fluid in the area of macroeconomics, 2009 Republicans appear more similar to their 1992 counterparts in their embrace of monetarist and supply side views. By contrast, 2009 Democrats appear increasingly skeptical of supply side propositions and more supportive of activist fiscal policy. The most enduring divisive issues between Republicans and Democrats involve the distribution of income and regulation. Immigration also appears to be an issue which finds Republicans and Democrats on the opposite side of the fence.

METHODOLOGY, SAMPLE, AND MEASURES OF CONSENSUS

The methodology employed in this paper relies on work originally done by Kearl et al. (1979), continued by Alston, Kearl and Vaughan (1992), and Fuller and Geide-Stevenson (2003), who study consensus among economists on a number of economics propositions. This methodology has also been used to survey economists in different countries, economists in different fields, as well as non-economists (e.g. Frey et. al, 1984, Ricketts and Shoesmith, 1992, Whaples, 2005). These studies ask participants to indicate whether they agree, agree with proviso or disagree (or a similar scale) with a given set of economic propositions in the areas of microeconomics, macroeconomics, income distribution, and international economics. Our three surveys of delegates are based on the original set of propositions developed by Kearl et. al. (1979), primarily positive statements that reflect basic concepts covered in standard introductory economics textbooks. We also include several normative statements that reflect fundamental values which often shape debates concerning economic policy. The current survey contains 42 propositions, of which 37 are identical to propositions in the 2000 survey while 23 are identical to propositions in the 1992 survey.

Fuller, Alston, and Vaughan (1995) conducted the first survey of party delegates followed by a second survey by Fuller and Geide-Stevenson (2007). In each study, surveys were mailed to a random sample of 1,000 – 1,300 delegates from each party. In the current sample, 1,200 Democratic delegates and 1,300 Republican delegates were mailed surveys in the Spring of 2009. Difficulties in obtaining the Republican delegate list caused a slight delay in the date at which surveys were mailed to Republican delegates. Response rates are 10.6% for Democrats and 14.4% for Republicans, lower than the respective response rates of 17.5% and 15.8% for the 2000 survey. These response rates are substantially lower than the response rates of around 40% from other convention delegate studies (Herrera, 1992). We can only speculate that this may be

due to the more technical nature of our survey instrument, survey fatigue, or some other combination of factors.

Our empirical analysis utilizes two measures of consensus. Following the basic methodology of Kearl, et. al., our first measure of consensus is the relative entropy index, ε, which is calculated based on the probabilities, p_i of each possible outcome, $i = 1,\ldots,n$. For each economic proposition there are four possible outcomes, agree, agree with proviso, disagree or no response. Given the observed relative frequencies, p_i , the entropy index is constructed as $E(p_i) = \sum_{i=1}^{4} -p_i \log_2 p_i$. The relative entropy index, ε, for each proposition is calculated by dividing the entropy measure $E(p_i)$ by the maximum possible entropy which occurs when responses are equally distributed across all possible response options (i.e. p = 0.25). In short, the relative entropy index is defined as $\varepsilon = E(p_i)/(\text{maximum possible entropy})$. Given this definition, relative entropy ε can take on values between 0 and 1 where $\varepsilon = 0$ when all respondents choose the same response, that is, complete consensus. A relative entropy index of $\varepsilon = 1$ indicates all responses are equally likely, that is, no consensus. Thus, the lower the entropy index, the higher the degree of consensus on a specific proposition. As Fuller *et al*. (1995) indicate, the relative entropy index is nonlinear, as small changes in the distribution of responses result in large changes in entropy. For example, a response pattern of 70-15-10-5 (in percent) generates a relative entropy index of 0.66 while a response pattern of 60-20-15-5 results in an entropy index of 0.77. Following Fuller and Geide-Stevenson (2007), we define $\varepsilon \leq 0.8$ to indicate consensus and construct a conditional measure of broad agreement. This measure is useful because it indicates the direction of opinion. We first add the frequency of those who "generally agree" to those who "agree with provisos". We then divide by the total number of responses less the frequency of those who returned "no response" to the proposition. In this way, we split respondents' opinions into "broadly agree" or "disagree". This second measure is taken to indicate consensus when at least 67% of respondents either broadly agree or disagree. When both the relative entropy index and the conditional percentage indicate consensus, we conclude "strong consensus". When only one of our measures indicates consensus, we conclude "consensus", and when neither measure indicates consensus, we conclude "no consensus".

EMPIRICAL RESULTS

Relative frequencies of responses for all three surveys are reported in Table 1 along with relative entropy indices, conditional percentages of agreement/ disagreement and conclusions of consensus. In addition, Table 1 also includes the p-values for the standard chi-square test of independence for 2000 and 2008 Republican and Democratic delegations, the 2000 and 2008 Republican delegations, and the 2000 and 2008 Democratic delegations. We use this to test the null hypothesis that the distribution of responses within a party is independent of when the survey was conducted. This test helps determine if response patterns on specific propositions have changed significantly over time. Since the chi-square test of independence is only useful when each response category is observed in sufficient numbers and the proportion of 'no response' is generally low or zero in our survey, we exclude the 'no response' category when

performing chi-square tests. We use the 5% level of significance ($p \leq 0.05$) in order to reject the null hypothesis. Hence, rejection of the null hypothesis implies a high likelihood that the distribution of responses has changed over time.

For example, 46.8% of Republicans agreed with proposition #2 in 2000, while 70.4% agreed with this proposition in 2009. The p-value comparing the response distribution is 0 indicating that the response pattern has changed with certainty. By contrast, 31.4% and 34.4% of Democrats agree with proposition #2 in 2000 and 2009. The p-value of 0.354 indicates that the null hypothesis cannot be rejected, the response distribution is likely identical over time. Due to the complexity of the table, we do not report tests of independence involving the 1992 national delegations, referring to them only as warranted.

Table 1
Distribution of Responses, Measures of Consensus

		Republican			*Democrat*			Chi-Square P-values
		1992	2000	2008	*1992*	*2000*	*2008*	
Microeconomic Propositions								
1. An economy that operates below potential GDP has a self- correcting mechanism that will eventually return it to potential real GDP.	A[1]	70.2	32.9	48.6	*13.2*	*15.4*	*6.4*	**R00-D00**
	A/P	16.2	34.2	29.1	*13.2*	*36.0*	*31.2*	*p= 0.00*
	D	12.4	9.5	18.4	*72.1*	*32.6*	*53.6*	**R08-D08**
	NR	1.2	23.4	3.9	*1.5*	*16.0*	*8.8*	*p= 0.00*
	ε	.60	.94	.83	*.60*	*.95*	*.78*	**R00-R08**
	AG/DG	.97/.13	.88/.12	.81/.19	*.27/.73*	*.61/.39*	*.41/.59*	*p= 0.09*
	Concl.	Str.	Cons.	Cons.	*Str.*	*Cons.*	*Cons.*	**D00-D08** *p= 0.00*
2. There is a natural rate of unemployment to which the economy tend in the long run	A	80.7	46.8	70.4	*43.2*	*31.4*	*34.4*	**R00-D00**
	A/P	13.5	36.2	19.6	*29.3*	*26.3*	*29.6*	*p*= 0.00
	D	4.2	13.9	6.7	*26.4*	*36.6*	*28.0*	**R08-D08**
	NR		3.2	3.3		*5.7*	*8.0*	*p*= 0.00
	ε	.47	.80	.62	*.81*	*.90*	*.93*	**R00-R08**
	AG/DG	.96/.04	.86/.14	.93/.07	*.73/.27*	*.61/.39*	*.70/.30*	*p*= 0.00
	Concl.	Str.	Str.	Str.	*Cons.*	*None*	*Cons.*	**D00-D08** *p*= 0.35
3. In the short run, a reduction in unemployment causes the rate of inflation to increase*	A	58.3	17.7	16.2	*23.2*	*14.3*	*12.8*	**R00-D00**
	A/P	27.6	28.5	25.1	*35.7*	*30.3*	*21.6*	*p*= 0.67
	D	12.1	47.5	53.6	*40.0*	*49.7*	*58.4*	**R08-D08**
	NR		6.3	5.0		*5.7*	*7.2*	*p*= 0.54
	ε	.73	.86	.81	*.81*	*.83*	*.79*	**R00-R08**
	AG/DG	.88/.12	.49/.51	.44/.56	*.73/.27*	*.47/.53*	*.37/.62*	*p*= 0.58
	Concl.	Str.	None	None	*Cons.*	*None*	*Cons.*	**D00-D08** *p*= 0.20
4. Changes in aggregate demand affect real GDP in the short run but not in the long run*	A	-	18.4	22.9	-	*18.3*	*12.0*	**R00-D00**
	A/P	-	37.3	29.6	-	*35.4*	*24.0*	*p*= 0.83
	D	-	26.0	37.4	-	*29.1*	*40.8*	**R08-D08**
	NR	-	18.4	10.1	-	*17.1*	*23.2*	*p*= 0.11
	ε	-	.97	.94	-	*.97*	*.94*	**R00-R08**

Table 1
Distribution of Responses, Measures of Consensus

		Republican			*Democrat*			Chi-Square P-values
		1992	2000	2008	*1992*	*2000*	*2008*	
	AG/DG	-	.68/.32	.58/.42	-	*.65/.35*	*.47/.53*	p= 0.08
	Concl.	-	Cons.	None	-	*None*	*None*	**D00-D08** p= 0.02
5. Inflation is caused primarily by too much growth in the money supply.	A	47.9	34.8	60.3	*25.0*	*24.0*	*22.4*	**R00-D00** p= .11
	A/P	23.5	27.9	22.9	*42.9*	*33.1*	*28.0*	
	D	26.6	32.9	15.6	*30.0*	*36.6*	*40.8*	**R08-D08** p= .00
	NR		4.4	1.1		*6.3*	*8.8*	
	ε	.81	.89	.71	*.83*	*.90*	*.92*	**R00-R08** p= .00
	AG/DG	.73/.27	.66/.34	.84/.16	*.69/.31*	*.61/.39*	*.55/.45*	
	Concl.	Cons.	Str.	Str.	*Cons.*	*None*	*None*	**D00-D08** p= .61
6. The Federal Reserve should focus on a low rate of inflation rather than other possible goals such as employment, or economic growth.	A	-	43.0	41.3	-	*16.0*	*9.6*	**R00-D00** *p= 0.00*
	A/P	-	31.7	34.1	-	*26.9*	*22.4*	
	D	-	22.2	21.8	-	*54.9*	*64.8*	**R08-D08** *p= 0.00*
	NR	-	3.2	2.8	-	*2.3*	*3.2*	
	ε	-	.84	.84	-	*.77*	*.69*	**R00-R08** *p= 0.90*
	AG/DG	-	.77/.23	.78/.22	-	*.44/.56*	*.33/.67*	
	Concl.	-	Cons.	Cons.	-	*Cons.*	*Str.*	**D00-D08** *p= 0.13*
7. Management of the business cycle should be left to the Federal Reserve; activist fiscal policy should be avoided	A	34.5	50.0	39.1	*10.7*	*25.1*	*9.6*	**R00-D00** *p= 0.00*
	A/P	31.0	31.7	28.5	*26.8*	*29.7*	*32.8*	
	D	32.4	15.8	25.7	*60.0*	*36.6*	*52.0*	**R08-D08** *p= 0.00*
	NR		2.5	6.7		*8.6*	*5.6*	
	ε	.85	.79	.91	*.71*	*.93*	*.79*	**R00-R08** *p= 0.04*
	AG/DG	.67/.33	.84/.16	.72/.28	*.38/.62*	*.60/.40*	*.45/.55*	
	Concl.	Cons.	Str.	Cons.	*Cons.*	*None*	*Cons.*	**D00-D08** *p= 0.00*
8. Increasing the regulatory power of the Federal Reserve will improve the functioning of financial markets.	A	-	-	6.7	-	-	*53.6*	**R08-D08** *p= 0.00*
	A/P	-	-	30.2	-	-	*26.4*	
	D	-	-	62.0	-	-	*18.4*	
	NR	-	-	1.1	-	-	*1.6*	
	ε	-		.64	-		*.77*	
	AG/DG	-		.37/.63	-		*.81/.19*	
	Concl.	-		Cons.	-		*Str.*	
9. Fiscal policy has a significant stimulative impact on a less than fully employed economy.	A	70.3	50.6	34.1	*64.3*	*37.7*	*56.8*	**R00-D00** *p= 0.00*
	A/P	16.9	38.6	31.3	*16.8*	*33.7*	*29.6*	
	D	10.7	8.9	31.8	*18.2*	*22.3*	*8.8*	**R08-D08** *p= 0.00*
	NR		1.9	2.7		*6.3*	*4.8*	
	ε	.63	.72	.90	*.67*	*.90*	*.75*	**R00-R08** *p= 0.00*
	AG/DG	.89/.11	.91/.09	.76/.24	*.82/.18*	*.76/.24*	*.91/.09*	

Table 1
Distribution of Responses, Measures of Consensus

		Republican			*Democrat*			Chi-Square P-values
		1992	2000	2008	*1992*	*2000*	*2008*	
	Concl.	Str.	Str.	Cons.	*Str.*	*Cons.*	*Str.*	**D00-D08** *p= 0.00*
10. A large federal budget deficit has an adverse effect on the economy.	A	89.3	65.8	91.1	*86.1*	*72.6*	*38.4*	**R00-D00**
	A/P	5.9	22.2	7.8	*5.4*	*17.1*	*44.0*	*p= 0.32*
	D	3.1	12.0	1.1	*7.1*	*9.1*	*15.2*	**R08-D08**
	NR		0.0	0.0		*1.1*	*2.4*	*p= 0.00*
	ε	.32	.62	.24	*.39*	*.58*	*.80*	**R00-R08**
	AG/DG	.97/.03	.88/.12	.99/.01	*.93/.07*	*.91/.09*	*.84/.16*	*p= 0.00*
	Concl.	Str.	Str.	Str.	*Str.*	*Str.*	*Str.*	**D00-D08** *p= 0.00*
11. If the federal budget is to be balanced, it should be done over the course of the business cycle rather than yearly.	A	33.8	19.0	18.4	*34.3*	*18.3*	*22.4*	**R00-D00**
	A/P	27.6	31.0	23.5	*40.7*	*26.3*	*24.8*	*p= 0.86*
	D	36.9	46.8	54.7	*23.2*	*45.7*	*41.6*	**R08-D08**
	NR		3.2	3.4		*9.7*	*11.2*	*p= 0.25*
	ε	.84	.83	.779	*.83*	*.90*	*.93*	**R00-R08**
	AG/DG	.62/.38	.52/.48	.43/.57	*.76/.24*	*.40/.51*	*.53/.47*	*p= 0.25*
	Concl.	None	None	Cons.	*Cons.*	*None*	*None*	**D00-D08** *p= 0.62*
12. The level of government spending relative to GDP should be reduced (disregarding expenditures for stabilization).	A	80.3	62.0	86.6	*45.0*	*17.1*	*16.8*	**R00-D00**
	A/P	13.1	27.9	7.8	*29.3*	*30.9*	*34.4*	*p= 0.00*
	D	3.8	6.3	4.5	*22.5*	*44.0*	*41.6*	**R08-D08**
	NR		3.8	1.1		*8.0*	*7.2*	*p= 0.00*
	ε	.48	.69	.37	*.84*	*.87*	*.88*	**R00-R08**
	AG/DG	.96/.04	.93/.07	.95/.05	*.76/.24*	*.52/.48*	*.55/.45*	*p= 0.00*
	Concl.	Str.	Str.	Str.	*Cons.*	*None*	*None*	**D00-D08** *p= 0.83*
13. Appropriately designed fiscal policy can increase the long run rate of capital formation.	A	-	55.1	55.3	-	*49.1*	*54.4*	**R00-D00**
	A/P	-	32.9	29.1	-	*37.7*	*32.0*	*p= 0.22*
	D	-	4.4	11.8	-	*1.7*	*4.8*	**R08-D08**
	NR	-	7.6	4.5	-	*11.4*	*8.8*	*p= 0.17*
	ε	-	.74	.77	-	*.75*	*.76*	**R00-R08**
	AG/DG	-	.95/.05	.88/.12	-	*.98/.02*	*.95/.05*	*p= 0.08*
	Concl.	-	Str.	Str.	-	*Str.*	*Str.*	**D00-D08** *p= 0.19*
14. Lower marginal income tax rates reduce leisure and increase work effort.	A	33.5	18.4	37.4	*9.6*	*11.4*	*3.2*	**R00-D00**
	A/P	29.3	17.7	18.4	*25.0*	*14.9*	*12.0*	*p= 0.18*
	D	35.2	59.5	39.1	*63.9*	*65.1*	*73.6*	**R08-D08**
	NR		4.4	5.0		*8.6*	*11.2*	*p= 0.00*
	ε	.85	.77	.86	*.60*	*.74*	*.60*	**R00-R08**
	AG/DG	.64/.36	.38/.62	.59/.41	*.35/.65*	*.29/.71*	*.17/.83*	*p= 0.00*
	Concl.	None	Cons.	Cons.	*Cons.*	*Str.*	*Str.*	**D00-D08** *p= 0.03*

Table 1
Distribution of Responses, Measures of Consensus

		Republican			*Democrat*			Chi-Square P-values
		1992	2000	2008	*1992*	*2000*	*2008*	
15. Reducing the tax rate on income from capital gains would encourage investment and promote economic growth.	A	95.2	81.7	91.6	*24.6*	*24.6*	*18.4*	**R00-D00** *p= 0.00*
	A/P	3.1	15.2	5.0	*20.7*	*21.1*	*18.4*	
	D	0.7	1.3	2.2	*53.9*	*49.7*	*60.0*	**R08-D08** *p= 0.00*
	NR		1.9	1.1		*4.6*	*3.2*	
	ε	.17	.42	.26	*.75*	*.84*	*.75*	**R00-R08** *p= 0.01*
	AG/DG	.99/.01	.99/.01	.98/.02	*.46/.54*	*.48/.52*	*.38/.62*	
	Concl.	Str.	Str.	Str.	*Cons.*	*None*	*Cons.*	**D00-D08** *p= 0.23*
16. Managerial, information and other technological advances have significantly lessened the severity of or fundamentally eliminated the business cycle.	A	-	17.1	11.2	-	*15.4*	*10.4*	**R00-D00** *p= 0.83*
	A/P	-	32.9	24.0	-	*29.7*	*19.2*	
	D	-	41.8	58.7	-	*43.4*	*58.4*	**R08-D08** *p= 0.76*
	NR	-	8.2	6.1	-	*11.4*	*12.0*	
	ε	-	.89	.77	-	*.91*	*.81*	**R00-R08** *p= 0.01*
	AG/DG	-	.54/.46	.38/.63	-	*.51/.49*	*.34/.66*	
	Concl.	-	None	Cons.	-	*None*	*None*	**D00-D08** *p= 0.02*
17. The U.S. has entered a new industrial revolution in which higher rates of economic growth can be maintained without inflationary pressures.	A	-	41.1	16.2	-	*27.4*	*14.4*	**R00-D00** *p= 0.06*
	A/P	-	34.1	24.0	-	*38.9*	*23.2*	
	D	-	20.3	53.6	-	*25.1*	*54.4*	**R08-D08** *p= 0.92*
	NR	-	4.4	6.2	-	*8.6*	*8.0*	
	ε	-	.86	.82	-	*.92*	*.83*	**R00-R08** *p= 0.00*
	AG/DG	-	.79/.21	.43/.57	-	*.73/.27*	*.41/.59*	
	Concl.	-	Cons.	None	-	*Cons.*	*None*	**D00-D08** *p= 0.00*
International Economics Propositions								
18. Tariffs and import quotas usually reduce the general welfare of society.	A	62.4	41.1	53.1	*25.7*	*24.6*	*16.0*	**R00-D00** *p= 0.00*
	A/P	20.7	26.6	22.9	*28.6*	*25.1*	*20.0*	
	D	13.8	31.0	21.2	*40.4*	*45.1*	*58.4*	**R08-D08** *p= 0.00*
	NR		1.3	2.8		*5.2*	*5.6*	
	ε	.72	.82	.80	*.89*	*.87*	*.79*	**R00-R08** *p= 0.05*
	AG/DG	.86/.14	.69/.31	.78/.22	*.57/.43*	*.52/.48*	*.38/.62*	
	Concl.	Str.	Cons.	Str.	*None*	*None*	*Cons.*	**D00-D08** *p= 0.05*
19. Flexible and floating exchange rates offer an effective international monetary arrangement.	A	52.4	50.0	64.2	*35.4*	*38.9*	*24.8*	**R00-D00** *p= 0.10*
	A/P	37.2	34.8	26.8	*47.5*	*42.9*	*42.4*	
	D	4.8	7.6	4.5	*13.6*	*10.9*	*16.8*	**R08-D08** *p= 0.00*
	NR		7.6	4.5		*7.3*	*16.0*	
	ε	.73	.80	.66	*.80*	*.84.*	*.94*	**R00-R08** *p= 0.44*
	AG/DG	.95/.05	.92/.08	.95/.05	*.86/.14*	*.88/.12*	*.80/.20*	

Table 1
Distribution of Responses, Measures of Consensus

		Republican			*Democrat*			Chi-Square P-values
		1992	2000	2008	*1992*	*2000*	*2008*	
	Concl.	Str.	Str.	Str.	*Str.*	*Cons.*	*Cons.*	**D00-D08** *p= 0.06*
20. Increasing globalization of the economy, helped by the WTO, threatens nat'l sovereignty in the areas of environmental and labor standards.	A	-	30.4	48.0	-	*30.9*	*28.8*	**R00-D00**
	A/P	-	27.2	21.8	-	*21.1*	*19.2*	*p= 0.43*
	D	-	40.5	29.1	-	*45.1*	*48.0*	**R08-D08**
	NR	-	1.9	1.1	-	*2.9*	*4.0*	*p= 0.00*
	ε	-	.84	.79	-	*.83*	*.83*	**R00-R08**
	AG/DG	-	.59/.41	.71/.29	-	*.54/.46*	*.50/.50*	*p= 0.01*
	Concl.	-	None	Cons.	-	*None*	*None*	**D00-D08** *p= 0.84*
21. Easing restrictions on immigration will ensure long run economic growth.	A	-	-	8.4	-	-	*24.0*	
	A/P	-	-	14.0	-	-	*33.6*	**R08-D08**
	D	-	-	76.5	-	-	*39.2*	*p= 0.00*
	NR	-	-	1.1	-	-	*3.2*	
	ε	-		.53	-		*.86*	
	AG/DG	-		.23/.77	-		*.60/.41*	
	Concl.	-		Str.	-		*None*	
22. Easing restrictions on immigration will depress the average wage rate in the United States.	A	-	-	43.0	-	-	*20.8*	
	A/P	-	-	19.6	-	-	*13.6*	**R08-D08**
	D	-	-	34.6	-	-	*63.2*	*p= 0.00*
	NR	-	-	2.8	-	-	*2.4*	
	ε	-		.83	-		*.71*	
	AG/DG	-		.64/.36	-		*.35/.65*	
	Concl.	-		None	-		*Cons.*	
23. Large balance of trade deficits have adverse effects on the economy.	A	80.7	50.6	67.6	*82.5*	*54.3*	*69.6*	**R00-D00**
	A/P	12.1	26.6	17.9	*9.6*	*23.4*	*18.4*	*p=0.72*
	D	6.2	17.1	12.3	*5.7*	*15.4*	*7.2*	**R08-D08**
	NR		5.7	2.2		*6.9*	*4.8*	*p= 0.39*
	ε	.47	.84	.66	*.45*	*.83*	*.65*	**R00-R08**
	AG/DG	.94/.06	.82/.18	.87/.13	*.94/.06*	*.83/.17*	*.92/.08*	*p= 0.02*
	Concl.	Str.	Cons.	Str.	*Str.*	*Cons.*	*Str.*	**D00-D08** *p= 0.02*
24. The U.S. trade deficit is primarily due to non-tariff trade barriers erected by other nations.	A	-	20.9	11.7	-	*21.7*	*15.2*	**R00-D00**
	A/P	-	28.5	19.5	-	*22.9*	*16.0*	*p= 0.41*
	D	-	41.1	65.4	-	*47.4*	*63.2*	**R08-D08**
	NR	-	9.5	3.4	-	*8.0*	*5.6*	*p= 0.55*
	ε	-	.92	.69	-	*.88*	*.74*	**R00-R08**
	AG/DG	-	.55/.45	.32/.68	-	*.48/.52*	*.33/.67*	*p= 0.00*
	Concl.	-	None	Str.	-	*None*	*Str.*	**D00-D08** *p= 0.00*
25. The economic benefits of an	A	-	27.2	24.6	-	*10.9*	*6.4*	**R00-D00**
	A/P	-	17.1	24.6	-	*24.0*	*15.2*	*p= 0.00*

Table 1
Distribution of Responses, Measures of Consensus

		Republican			*Democrat*			Chi-Square P-values
		1992	2000	2008	*1992*	*2000*	*2008*	
expanding world population outweigh the economic costs.	D	-	50.0	49.1	-	*61.1*	*73.6*	**R08-D08**
	NR	-	5.7	1.7	-	*4.0*	*4.8*	*p= 0.00*
	ε	-	.84	.80	-	*.73*	*.60*	**R00-R08**
	AG/DG	-	.47/.53	.50/.50	-	*.36/.64*	*.23/.77*	*p= 0.31*
	Concl.	-	None	Cons.	-	*Cons.*	*Str.*	**D00-D08**
								p= 0.05
26. Some restrictions on the free flow of financial capital are essential to ensure the stability and soundness of the international financial system.	A	-	20.9	19.6	-	*42.3*	*52.0*	**R00-D00**
	A/P	-	33.5	41.9	-	*40.6*	*30.4*	*p= 0.00*
	D	-	37.3	35.7	-	*11.4*	*8.8*	**R08-D08**
	NR	-	8.3	2.8	-	*5.7*	*8.8*	*p= 0.00*
	ε	-	.91	.83	-	*.82*	*.81*	**R00-R08**
	AG/DG	-	.59/.41	.63/.37	-	*.88/.12*	*.90/.10*	*p= 0.50*
	Concl.	-	None	None	-	*Cons.*	*Cons.*	**D00-D08**
								p= 0.14
Distribution of Income and Wealth Propositions								
27. The distribution of income in the U.S. should be more equal	A	10.0	6.3	2.2	*76.4*	*62.9*	*59.2*	**R00-D00**
	A/P	19.7	10.8	11.2	*13.6*	*30.9*	*29.6*	*p= 0.00*
	D	69.7	82.3	84.9	*8.9*	*6.3*	*9.6*	**R08-D08**
	NR		0.6	1.7		*0.0*	*1.6*	*p= 0.00*
	ε	.60	.44	.39	*.53*	*.60*	*.69*	**R00-R08**
	AG/DG	.30/.70	.17/.83	.14/.86	*.91/.09*	*.94/.06*	*.90/.10*	*p= 0.18*
	Concl.	Str.	Str.	Str.	*Str.*	*Str.*	*Str.*	**D00-D08**
								p= 0.54
28. The increasing inequality in the distribution of income in the U.S. is due primarily to the benefits and pressures of a global economy.	A	-	10.1	11.7	-	*12.6*	*6.4*	**R00-D00**
	A/P	-	19.6	14.5	-	*16.0*	*12.8*	*p= 0.58*
	D	-	65.3	71.5	-	*68.0*	*76.0*	**R08-D08**
	NR	-	5.0	2.2	-	*3.4*	*4.8*	*p= 0.28*
	ε	-	.70	.62	-	*.67*	*.57*	**R00-R08**
	AG/DG	-	.31/.69	.27/.73	-	*.30/.70*	*.20/.80*	*p= 0.40*
	Concl.	-	Str.	Str.	-	*Str.*	*Str.*	**D00-D08**
								p= 0.14
29. The redistribution of income within the U.S. is a legitimate role for government.	A	3.8	1.9	1.7	*55.4*	*40.0*	*49.6*	**R00-D00**
	A/P	7.9	5.7	3.4	*17.4*	*36.0*	*24.8*	*p= 0.00*
	D	86.6	91.1	93.8	*25.4*	*22.3*	*22.4*	**R08-D08**
	NR		1.3	1.1		*1.7*	*3.2*	*p= 0.00*
	ε	.38	.27	.21	*.76*	*.82*	*.82*	**R00-R08**
	AG/DG	.12/.88	.08/.92	.05/.95	*.74/.26*	*.77/.23*	*.77/.23*	*p= 0.57*
	Concl.	Str	Str.	Str.	*Str.*	*Cons.*	*Cons.*	**D00-D08**
								p= 0.11

Table 1
Distribution of Responses, Measures of Consensus

		Republican			*Democrat*			Chi-Square P-values
		1992	2000	2008	*1992*	*2000*	*2008*	
30. The distribution of income and wealth in the U.S. has little if any impact on the overall rate of economic growth and stability.	A	-	31.7	34.6	-	*6.3*	*4.8*	**R00-D00**
	A/P	-	26.0	16.2	-	*15.4*	*12.8*	*p= 0.00*
	D	-	39.2	46.4	-	*75.4*	*80.0*	**R08-D08**
	NR	-	3.2	2.8	-	*2.9*	*2.4*	*p= 0.00*
	ε	-	.86	.81	-	*.56*	*.49*	**R00-R08**
	AG/DG	-	.59/.41	.52/.48	-	*.22/.78*	*.18/.82*	*p= 0.08*
	Concl.	-	None	None	-	*Str.*	*Str.*	**D00-D08** *p= 0.66*
31. Minimum wages increase unemployment among young and unskilled workers.	A	64.5	45.6	65.4	*14.6*	*8.0*	*8.8*	**R00-D00**
	A/P	13.1	26.0	19.5	*10.7*	*8.0*	*2.4*	*p= 0.00*
	D	20.7	27.2	14.5	*73.2*	*82.3*	*88.8*	**R08-D08**
	NR		1.3	0.56		*1.7*	*0.0*	*p= 0.00*
	ε	.68	.81	.65	*.58*	*.46*	*.29*	**R00-R08**
	AG/DG	.79/.21	.72/.28	.85/.15	*.26/.74*	*.16/.84*	*.11/.89*	*p= 0.00*
	Concl.	Str.	Cons.	Str.	*Str.*	*Str.*	*Str.*	**D00-D08** *p= 0.11*
32. There are few compensation and promotion gaps between men and women that cannot be explained by productivity and/or career choices.	A	33.5	37.3	47.5	*5.7*	*10.9*	*12.0*	**R00-D00**
	A/P	25.5	24.7	25.7	*7.5*	*9.7*	*5.6*	*p= 0.00*
	D	40.0	36.1	26.8	*84.3*	*78.3*	*80.8*	**R08-D08**
	NR		1.9	0.0		*1.1*	*1.6*	*p= 0.00*
	ε	.81	.83	.76	*.43*	*.51*	*.47*	**R00-R08**
	AG/DG	.60/.40	.63/.37	.73/.27	*.13/.87*	*.21/.79*	*.18/.82*	*p= 0.11*
	Concl.	None	None	Str.	*Str.*	*Str.*	*Str.*	**D00-D08** *p= 0.43*
33. Welfare reforms which place time limits on public assistance have increased the general well-being of society.	A	-	77.9	79.3	-	*21.7*	*23.2*	**R00-D00**
	A/P	-	17.1	12.8	-	*40.6*	*35.2*	*p= 0.00*
	D	-	5.1	7.3	-	*37.1*	*40.8*	**R08-D08**
	NR	-	0.0	0.56	-	*0.6*	*0.8*	*p= 0.00*
	ε	-	.47	.48	-	*.79*	*.80*	**R00-R08**
	AG/DG	-	.95/.05	.93/.07	-	*.63/.37*	*.59/.41*	*p= 0.43*
	Concl.	-	Str.	Str.	-	*Cons.*	*Cons.*	**D00-D08** *p= 0.65*
34. The persistence of poverty is due more to a breakdown of the family unit than to a general lack of economic opportunity.	A	67.9	62.0	69.8	*8.6*	*17.7*	*5.6*	**R00-D00**
	A/P	15.9	24.1	18.4	*9.6*	*16.0*	*16.8*	*p= 0.00*
	D	15.5	13.3	11.2	*80.7*	*64.6*	*73.6*	**R08-D08**
	NR		0.6	1.1		*1.7*	*4.0*	*p= 0.00*
	ε	.64	.68	.62	*.47*	*.69*	*.59*	**R00-R08**
	AG/DG	.84/.16	.87/.13	.89/.11	*.18/.82*	*.34/.66*	*.23/.77*	*p= 0.33*
	Concl.	Str.	Str.	Str.	*Str.*	*Cons.*	*Str.*	**D00-D08** *p= 0.01*
35. The Earned	A	-	30.4	24.0	-	*59.4*	*47.2*	**R00-D00**

Table 1
Distribution of Responses, Measures of Consensus

		Republican			*Democrat*			Chi-Square P-values
		1992	2000	2008	*1992*	*2000*	*2008*	
Income Tax Credit program should be expanded.	A/P	-	22.8	16.8	-	*25.1*	*24.8*	*p= 0.00*
	D	-	42.4	55.3	-	*10.3*	*18.4*	**R08-D08**
	NR	-	4.4	3.9	-	*5.1*	*9.6*	*p= 0.00*
	ε	-	.87	.79	-	*.75*	*.89*	**R00-R08**
	AG/DG	-	.56/.44	.42/.58	-	*.89/.11*	*.80/.20*	*p= 0.06*
	Concl.	-	None	Cons.	-	*Str.*	*Cons.*	**D00-D08**
								p= 0.07
Microeconomics Propositions								
36. Antitrust laws should be enforced vigorously to reduce monopoly power from its current level.	A	34.5	21.5	27.4	*75.7*	*58.3*	*72.8*	**R00-D00**
	A/P	40.0	36.1	38.0	*17.1*	*30.3*	*22.4*	*p= 0.00*
	D	23.5	41.1	33.5	*6.1*	*10.3*	*3.2*	**R08-D08**
	NR		1.3	1.1		*1.1*	*1.6*	*p= 0.00*
	ε	.83	.81	.82	*.53*	*.69*	*.54*	**R00-R08**
	AG/DG	.76/.24	.58/.42	.66/.34	*.94/.06*	*.90/.10*	*.97/.03*	*p= 0.28*
	Concl.	Cons.	None	None	*Str.*	*Str.*	*Str.*	**D00-D08**
								p= 0.01
37. Pollution taxes or marketable pollution permits are a more economically efficient approach to pollution control than emission standards.	A	32.4	27.2	11.2	*24.6*	*14.9*	*20.8*	**R00-D00**
	A/P	37.2	24.1	14.5	*23.9*	*19.4*	*24.8*	*p= 0.00*
	D	27.9	44.9	72.6	*47.9*	*62.9*	*50.4*	**R08-D08**
	NR		3.8	1.7		*2.9*	*4.0*	*p= 0.00*
	ε	.85	.85	.69	*.84*	*.72*	*.83*	**R00-R08**
	AG/DG	.71/.29	.53/.47	.26/.74	*.53/.47*	*.35/.65*	*.48/.52*	*p= 0.00*
	Concl.	Cons.	None	Str.	*None*	*Cons.*	*None*	**D00-D08**
								p= 0.11
38. Higher taxes on fossil fuels will encourage firms to develop alternative energies that reduce carbon emissions	A	-	-	12.9	-	-	*58.4*	
	A/P	-	-	22.9	-	-	*26.4*	**R08-D08**
	D	-	-	60.9	-	-	*12.8*	*p= 0.00*
	NR	-	-	3.3	-	-	*2.4*	
	ε	-		.73	-		*.73*	
	AG/DG	-		.37/.63	-		*.87/.13*	
	Concl.	-		Cons.	-		*Str.*	
39. Reducing the regulatory power of the Environmental Protection Agency (EPA) would improve the economic efficiency of the U.S. economy.	A	56.2	51.3	65.4	*6.1*	*4.6*	*5.6*	**R00-D00**
	A/P	23.5	27.9	21.2	*8.9*	*12.0*	*8.0*	*p= 0.00*
	D	19.7	20.3	12.3	*83.6*	*80.0*	*84.0*	**R08-D08**
	NR		0.6	1.1		*3.4*	*2.4*	*p= 0.00*
	ε	.73	.76	.66	*.43*	*.50*	*.43*	**R00-R08**
	AG/DG	.80/.20	.80/.20	.88/.12	*.15/.85*	*.17/.83*	*.14/.86*	*p= 0.02*
	Concl.	Str.	Str.	Str.	*Str.*	*Str.*	*Str.*	**D00-D08**
								p= 0.49
40. Economic evidence suggests	A	19.3	13.9	11.7	*15.0*	*17.7*	*15.2*	**R00-D00**
	A/P	39.3	16.5	12.3	*39.6*	*16.6*	*15.2*	*p= 0.52*

Table 1
Distribution of Responses, Measures of Consensus

		Republican			*Democrat*			Chi-Square P-values
		1992	2000	2008	*1992*	*2000*	*2008*	
there are too many resources in American agriculture.	D	38.3	63.3	69.3	*41.4*	*56.6*	*55.2*	**R08-D08**
	NR		6.3	6.7		*9.1*	*14.4*	*p= 0.22*
	ε	.84	.75	.68	*.82*	*.83*	*.85*	**R00-R08**
	AG/DG	.60/.40	.32/.86	.26/.74	*.57/.43*	*.38/.62*	*.36/.64*	*p= 0.41*
	Concl.	None	Str.	Str.	*None*	*None*	*None*	**D00-D08**
								p= 0.92
41. Employer-provided health insurance reduces the efficiency of the labor market by reducing labor mobility.	A	-	-	18.4	-	-	*25.6*	
	A/P	-	-	19.0	-	-	*18.4*	**R08-D08**
	D	-	-	59.8	-	-	*55.2*	*p= 0.37*
	NR	-	-	2.8	-	-	*0.8*	
	ε	-		.75	-		*.74*	
	AG/DG	-		.39/.61	-		*.44/.56*	
	Concl.	-		Cons.	-		*Cons.*	
42. The competitive model is generally more useful for understanding the U.S. economy than are models of imperfect competition and other game theoretic models.	A	-	54.4	55.9	-	*29.7*	*20.8*	**R00-D00**
	A/P	-	26.0	26.3	-	*30.3*	*28.0*	*p= 0.00*
	D	-	5.1	6.1	-	*20.0*	*30.4*	**R08-D08**
	NR	-	14.6	11.7	-	*20.0*	*20.8*	*p= 0.00*
	ε	-	.80	.79	-	*.99*	*.99*	**R00-R08**
	AG/DG	-	.94/.06	.93/.07	-	*.75/.25*	*.62/.38*	*p= 0.94*
	Concl.	-	Str.	Str.	-	*Cons.*	*Cons.*	**D00-D08**
								p= 0.06

1: The possible responses are: A = Mainly agree, A/P = Agree with provisos, D = Disagree, NR = No response.
2: Records the frequencies of responses from the 2000 and 2009 sample.
3. Conditional percentages of broad agreement: AG = (A+A/P)/ (A+A/P+D) and disagreement DG = D/(A+A/P+D).
4: Columns 6, 7, 8 and 9 report the entropy index ε; the conditional percentage of broad agreement (AG) or disagreement (DG), and the level of consensus (strong, consensus or no consensus.
5: p-value for the chi-square test of identical distributions of responses between two groups, e.g., Republicans (R) and democrats (D).
6: Strong consensus: $\varepsilon \leq 0.8$ and AG or DG $\geq 67\%$.
7: Consensus: $\varepsilon \leq 0.8$ or AG or DG $\geq 67\%$.
8: No consensus: $\varepsilon < 0.8$ and AG or DG $< 67\%$.
9. "-": proposition was not included in that year's survey

CONSENSUS WITHIN THE DEMOCRATIC AND REPUBLICAN PARTIES OVER TIME

We define polarization to occur when opinions diverge towards poles of distribution. Thus, one indication of polarization is lower values of the relative entropy index as opinions migrate to either agreement or disagreement with a proposition. Said differently, a higher value

of the entropy index indicates more diversity of opinion. Table 2 displays the average relative entropy index for all propositions included in the surveys, the 23 propositions common to all surveys and the 37 propositions common to the 2000 and 2009survey only.

Table 2
Comparison of average entropy

	Average Relative Entropy		
	1992	**2000**	**2009**
Republicans	0.676	0.775	0.689
Democrats	0.679	0.775	0.746
Republicans (23)	0.653	0.739	0.641
Democrats (23)	0.668*	0.770*	0.730
Republicans (37)	…	0.770*	0.688*
Democrats (37)	…	0.785	0.745

At first glance, average entropy increased for both Democrats and Republicans between 1992 and 2000. Thus, there is the suggestion that both parties may have been more inclusive or diverse in opinion in 2000 than 1992. Comparing 2009 to 2000, we see a much steeper decline in the entropy index for Republicans than Democrats. When we limit our analysis to propositions that are common between surveys, the only statistically significant changes in mean entropy at the 5% significance level, indicated with an asterisk, are observed for Republican delegates between 2000 and 2009 (37 common propositions) and for Democratic delegates between 1992 and 2000 (23 common propositions).

We shed additional light on the convergence of opinion in each party by examining the incidence of strong consensus, consensus, and no consensus constructed from the entropy index and conditional percentages of agreement reported in Table 1. Table 3 summarizes the results for the 23 propositions common to all surveys, the 37 propositions common to the 2000 and 2009 surveys, and the 42 propositions of the 2009 survey.

In all three surveys, Republican Delegations report a higher incidence of strong consensus than the Democratic delegation. This result is invariant to the set of propositions. The overall sample proportions of strong consensus fell for Republicans from 1992 to 2000 at a 10% level of significance. However, the proportion of strong consensus among Republicans is higher in 2009 than in 2000 for the 37 common propositions at a 10% level of significance. For the 23 common propositions, the difference in the proportion of strong consensus is not statistically different among Republicans between 1992 and 2009. Among Democrats, the proportion of strong consensus for the entire set of propositions as well as the 23 common propositions falls

from 1992 to 2000 at a 10% level of significance. All other differences in the proportions of strong consensus among the Democratic delegations are insignificant.

Table 3
Comparison of Degree of Consensus

Sample	Strong Consensus	Consensus	No consensus
Republicans (2008)	16/23 22/37 23/42 (70%) (59%) (55%)	5/23 8/37 11/42 (22%) (22%) (26%)	2/23 7/37 8/42 (9%) (19%) (19%)
Democrats (2008)	10/23 16/37 18/42 (43%) (43%) (43%)	8/23 11/37 13/42 (35%) (30%) (31%)	5/23 10/37 11/42 (22%) (27%) (26%)
Republicans (2000)	12/23 16/37 (52%) (43%)	5/23 8/37 (22%) (22%)	6/23 13/37 (23%) (35%)
Democrats (2000)	7/23 11/37 (30%) (30%)	6/23 12/37 (26%) (32%)	10/23 14/37 (43%) (38%)
Republicans (1992)	15/23 (65%)	4/23 (17%)	4/23 (17%)
Democrats (1992)	12/23 (52%)	7/23 (30%)	4/23 (17%)

Taken together, the two measures of consensus suggest that Democrats became significantly more inclusive or diverse in economic opinion between 1992 and 2000. A slightly weaker conclusion follows for Republicans between 1992 and 2000. However, between 2000 and 2009, the data suggests that while Republicans became significantly less inclusive or diverse in economic opinion and returned to 1992 levels in some cases, the diversity of opinion among Democrats remained largely unchanged. These insights support the findings based on the average entropy measures.

The process of polarization suggests that opinions are fluid over time and the direction of change is towards a greater degree of certainty. Comparing the 2000 and 2009 samples, our data does suggest some migration of opinion in the last decade. For the 37 common distributions in 2000 and 2009, Democrats show statistically different response patterns at the 5% confidence level for 13 propositions, while Republicans show changed response patterns for 17 propositions. Thus, Republican delegates showed a slightly higher frequency of significant shifts in response patterns between 2000 and 2009.

The economic views of Democrats are most fluid in the area of macroeconomics where the distribution of responses has significantly changed for almost half of the propositions between 2000 and 2009. In several cases, however, the changes indicate greater uncertainty than certainty. For example, in 2000 the conditional rate of agreement with the concept of a self-correcting mechanism of the economy (#1) was 61% while in the 2009 sample agreement fell to 41%, a view more representative of the 1992 sample.[1] Reflecting, perhaps, the start of the Great Recession in 2008, Democrats now indicate no-consensus with the "new economy" proposition #17. In addition, the incidence of broad agreement with the proposition that short run fluctuations in aggregate demand have no long run impacts on real GDP (#4) has fallen from 65% in 2000 to 47% in 2009.

We do find macroeconomic propositions for which Democrats have a higher level of consensus in 2009. In terms of managing the business cycle, 2009 Democrats are significantly more likely to agree with the efficacy of fiscal policy (#9) and to disagree with relegating responsibility solely to the Federal Reserve Bank (#7) although only 55% disagree with this latter proposition. 2009 Democrats are also significantly more likely to disagree with the supply side proposition linking lower marginal income tax rates to increased work effort (#14).

Democrats in 2009 are significantly more likely to agree with the proposition that large balance of trade deficits have adverse effects on the economy (#23), and to disagree with the proposition that tariffs and import quotas reduce the general welfare of society (#18). There is some suggestion that the views of current Democrats more closely resembles opinions in 1992, when over 80% of the respondents agreed that trade deficits have an adverse effect on the economy. The common denominator between 1992 and 2009 is that in both years, the U.S. economy was in a slow recovery from a recent recession. Given this, current Democrats are significantly less likely to blame non-tariff trade barriers for the U.S. trade deficit (#24). Finally, Democrats are significantly more likely to disagree that the economic benefits of an expanding world population outweigh the economic costs (#25).

In the area of distributional and microeconomic propositions (#29 - #42), there are only two propositions for which the distribution of responses in 2009 shows a significant change. Democrats are more significantly likely to disagree that the persistence of poverty is due more to the breakdown of the family than to lack of economic opportunity (#34) and more likely to agree that antitrust laws should be vigorously enforced (#36).

Republican delegates' views on macroeconomic propositions appear to have changed to a greater degree than Democrats, showing a significant change from 2000 to 2009 in the response pattern for 60% of comparable propositions. 2009 delegates appear to express some stronger monetarist and supply side sentiments compared to their 2000 counterparts, more similar to the 1992 survey in some cases. Compared to 2000, current Republicans are significantly more likely to agree with the notion that the economy tends to a natural rate of unemployment in the long-run (#2), to agree with the proposition that large federal deficits have adverse effects on the economy (#10), to agree with the proposition that the level of government spending should be reduced relative to GDP (#12), and to agree that inflation is linked to the money supply (#5). They are also significantly more likely to disagree with the proposition that fiscal policy has a significant stimulative impact on a less than fully employed economy (#9). Propositions #14 and #15 that emphasize the incentive effects of taxes also generate significantly higher likelihoods of agreement.

Not all evidence points to increasing consensus in the area of macroeconomics. In 2009, the rate of broad agreement with the proposition that management of the business should be left to the Federal Reserve Bank (#7) was 72%, down from 84% in 2000. As with Democrats, 2009 Republicans now indicate no-consensus with the "new economy" proposition #17.

While free trade is embraced more strongly (#18), agreement that the WTO threatens sovereignty in the areas of labor and environmental standards (#20) is significantly higher. Compared to the 2000 delegation, 2009 Republicans are significantly more likely to agree that trade deficits have an adverse effect on the economy (#23). However, there are significantly

higher levels of disagreement that U.S. trade deficits are linked to non-tariff barriers of trade (#24).

For the distributional propositions, 2009 Republicans are significantly more likely to agree with the proposition that minimum wages increase unemployment among young and unskilled workers (#31). Significant shifts in opinion on microeconomics propositions are observed exclusively in the area of environmental policies. Current Republicans show significantly increased disagreement with the proposition that pollution taxes or permits are more efficient than emission standards (#37) and increased agreement with the proposition that reducing the regulatory power of the EPA will increase economic efficiency (#39). For propositions #31 and #39, the distribution of opinion in 2009 is not significantly different at a 5% level from 1992**.** Proposition #37 is striking in that 2009 Republicans now disagree more strongly than their 1992 counterparts.

CONSENSUS BETWEEN THE DEMOCRATIC AND REPUBLICAN PARTIES OVER TIME

To test whether the distribution of responses differs between the political parties, we again use a chi-square test of independence, rejecting the null-hypothesis at p-values of 5% or less. In 2000, we could not reject the null-hypothesis for 42% of all propositions at a 5% level of significance. In 2009, we could not reject the null-hypothesis for 26% of all propositions. This result suggests less similarity in the economic views of Republicans and Democrats between 2000 and 2009. Based on the 23 propositions common to all survey periods, Republican and Democratic response patterns are statistically similar at the 5% level for 13% in 1992, 35% in 2000, and 17% in 2009. This comparison suggests that partisan polarization on economic issues in 1992 was at least as strong as in the most recent survey period.

It may be the case, however, that the chi-square tests of independence overstate the degree of dissimilarity between the views of Republicans and Democrats. An additional criterion involves a comparison of the direction of conditional agreement/disagreement on each proposition. That is, if both parties indicate a majority either conditionally agree or disagree with a proposition, then there is an indication of common ground even if the distribution of responses is statistically different.[2] It is those propositions for which the parties differ in the direction of conditional agreement/disagreement and that have statistically significantly different distributions that polarization is the greatest. Looking at the 2009 survey only, we identify 16 of 42 propositions that satisfy both criteria. This leaves somewhat less that two thirds of the propositions for which there appears to be common ground.

These 16 propositions are relatively clustered in propositions concerning the regulation and the distribution of income and wealth. There seems to be little common ground concerning the normative propositions that the distribution of income should be more equal (#27), that the redistribution of income is a legitimate role for government (#29), or that the Earned Income Tax Credit program should be expanded (#35). Similarly, Democrats and Republicans show little commonality for the positive propositions that the minimum wage increases unemployment among young and unskilled workers (#31), that few compensation and promotion gaps among

men and women are unexplained by productivity and/or career choices (#32), or that the persistence of poverty is due more to a breakdown of the family unit than lack of economic opportunity (#34). There is also little commonality concerning the propositions regarding regulation. Democrats and Republicans are of opposite opinion about the likelihoods that increasing the regulatory power of the Federal Reserve improving the functioning of financial markets (#9) and reductions in the regulatory power of the EPA improving the economic efficiency of the U.S. economy (#39). We also find evidence that the views of Republicans and Democrats concerning the benefits of easing restrictions on immigration (#21, #22) are on the opposite side of the fence.

One of the few propositions that generated an identical response pattern by both parties in each survey period is proposition #23 stating that large trade deficits have adverse effects on the economy. In 1992, delegates from both parties showed strong agreement with this proposition. While both 2009 delegations agree that flexible and floating exchange rates are an effective international monetary arrangement (#19), Democrats now disagree with the proposition that tariffs and import quotas usually reduce the general welfare of society (#18), a significant change from 2000. A consensus of agreement among 2009 Republicans with proposition #20 may be linked to the highlighted role of the World Trade Organization, betraying increased scepticism toward supra-national governing bodies.

In the area of macroeconomics, there is agreement in both 2009 delegations with the normative proposition that government spending should be reduced relative to GDP (#12) although Republicans more strongly agree than Democrats. In addition, both parties show a strong consensus of agreement that large federal deficits have adverse effects on the economy (#10). There does appear to be some divergence in 2009 concerning macroeconomic policy, however. While the level of agreement among Democrats on the link between money supply and inflation (#5) seems to have declined over time, Republicans more strongly embrace this monetarist view. Furthermore, there appears to be a divergence of opinion between 2009 Republicans and Democrats over the normative proposition that the Federal Reserve Bank should focus only on a low rate of inflation (#6). In the area of fiscal policy, 2009 Republicans retain their agreement with supply propositions (#13, #14, #15) while the level of disagreement among Democrats appears to be increasing over time.

CONCLUSION

The data suggest an increase in the degree of polarization between Republicans and Democrats from 2000 and 2009. At the aggregate level, there seem to be two trends that impact the apparent divergence of opinion between 2009 Democrats and Republicans. The first is a greater degree of consensus or convergence of opinion from 2000 to 2009 in the Republican party. In some respects, the 2009 Republican delegation resembles the 1992 delegation with a stronger embrace of monetarist and supply side views than in 2000. The second trend is the lower degree of consensus among 2000 compared to 1992 Democrats that was not reversed in 2009. Reflecting, perhaps the onset on the Great Recession, Democratic opinions in 2009 shifted

towards a stronger embrace of the efficacy of fiscal policy and increased doubt about the ability of the economy to self-correct.

While it is tempting to emphasize the extent of current polarization, we note many areas of agreement between Republicans and Democrats. Of the 42 propositions in the 2009 survey, both parties appear to be on the same side of the fence for 26 propositions in the sense that a majority broadly agree or disagree. There are several propositions for which there is evidence of continued agreement over time. For example, both parties agree that the level of government spending relative to GDP should be reduced and that well designed fiscal policy can increase the rate of capital formation. There is also shared concern about the impacts of large balance of trade deficits and federal budget deficits. We suggest that agreement among the parties is due to the generality and broadness of these propositions and disagreement arises over the tools used to address these issues.

It is notable the extent to which the opinions of Republicans and Democrats continue to differ when it comes to issues concerning the distribution of income. Republicans and Democrats continue to be strongly on the opposite side of the fence over the normative propositions concerning equality in the distribution of income and the legitimacy of the role of government in redistribution income. It is possible that these durable normative values spill over into opinions about the positive propositions such as the impact of minimum wages on unemployment among young and unskilled workers and the persistence of poverty. As economists note, almost every change in public policy, macroeconomic or microeconomic, has distributional implications. Economists are also adept at identifying the winners and losers of changes in public policy. Unfortunately, only under strict assumptions can economists render conclusions about distributional changes on the social welfare function. Given the strong polarization in views about the distribution of income, it may serve as an economic wedge issue and a driver of political gridlock. Unfortunately, it is in regards to the costs and benefits of income redistribution that economists have the least to offer.

Finally we note the polarization that is apparent in the propositions involving regulation and/or the environment. This is evident in the strong diversity of opinion concerning the stronger regulations evolving for the financial industry as well the substitution impacts of higher taxes on fossil fuels. There is also persistent polarization over the efficiency effects of reducing the power of the Environmental Protection Agency. And while the comparative faith in the ability of regulation to improve market outcomes has long distinguished the liberal from the conservative view, we note the inclusion in the 2008 Republican Platform of the call for "reasonable regulation, basing it on sound science to achieve goals that are technically feasible…". One possibility is that this shows an increasing skepticism of academic research by the Republican Party. If so, this may be a call for academicians as a whole to reflect on whether our normative values drive our research outcomes or whether our research informs our normative values.

Endnotes

1. Interestingly, there is still broad agreement on the existence of a natural unemployment rate to which the economy tends in the long-run (#2). As developed by Milton Friedman (1968) and Edmund Phelps (1968), the basic argument is that fiscal policy can help reduce unemployment to, but not sustainably below, the unemployment rate consistent with long run aggregate supply.

1. Fuller and Geide-Stevenson (2003) report that 91.6% of economists in their survey broadly agree with proposition #37. Given that it is now more or less standard in introductory economics texts to discuss the comparative efficiency of marketable pollution permits and/or effluent taxes versus emissions standards, the apparent growing disconnect between politician (both Republican and Democrat) and economist is glaring in light of the Fuller and Geide-Stevenson (2007) finding that Democrats and Republicans are more likely to agree among themselves than to agree with economists.
2. A good example of this is proposition #2 concerning the natural rate of unemployment. The 2009 sample conditional rate of agreement among Republicans is 93% while for Democrats it is 70%. However, the chi-square test of independence is rejected at a 5% level of significance.

REFERENCES

Alston, R. M., J.R. Kearl and M. B. Vaughan. (1992). Is there a Consensus Among Economists in the 1990's? *American Economic Journal* 82(2):203-209.

Beinart, Peter. (2010). Why Washington's Tied Up in Knots. *Time* 175(8):20-24.

Frey, B. S., W.W. Pommerehne, F. Schneider and G. Gilbert. (1984). Consensus and Dissension Among Economists: An Empirical Inquiry."American Economic Review 74(5):986-993.

Friedman, Milton (1968). The Role of Monetary Policy. *American Economic Review* 58: 1-17.

Fuller, Dan A., R. M. Alston and M. B. Vaughan (1995). The Split Between Political Parties on Economic Issues: A Survey of Republicans, Democrats, and Economists. *Eastern Economic Journal* 21(2):227 – 238.

Fuller, Dan A. and Doris Geide-Stevenson (2003). "Consensus Among Economists: Revisited." *Journal of Economic Education* 34(4):369-387.

Fuller, Dan A. and Doris Geide-Stevenson (2007). A Survey of Republicans, Democrats and Economists. *Eastern Economic Journal* 33(1):81-94.

Glaeser, Edward L, Giacomo A. M. Ponzetto and Jesse M. Shapiro (2005). Strategic Extremism: Why Republicans and Democrats Divide on Religious Values. *Quarterly Journal of Economics* 120(4):1283-1330.

Herrera, Richard (1992). The Understanding of Ideological Labels by Political Elites: A Research Note. *Western Political Quarterly*, 45(4):1021-1035.

Kearl, J. R., Clayne L. Pope, Gordon C. Whiting and Larry T. Wimmer (1979). A Confusion of Economists.. *American Economic Review*, 69(2):28-37.

Layman, Geoffrey C. (1999). "Culture Wars" in the American Party System: Religious and Cultural Change among Partisan Activists Since 1972. *American Politics Research* 27:89-121.

Martin, Jurek (2010). A Sad Descent into Vitriolic Polarisation. *Financial Times* (May 5):8

Miller, Warren E. and M. Kent Jennings (1992). *Parties in Transition: a Longitudinal Study of Party Elites and Party Supporters.* New York: Russell Sage Foundation.

Phelps, Edmund (1968). Money-Wage Dynamics and Labor-Market Equilibrium. *Journal of Political Economy* 76: 678-711.

Ricketts, M. and E. Shoesmith (1992). British Economic Opinion: Positive Science of Normative Judgement? *American Economic Review* 82(2):210-215.

The 2008 Democratic National Platform. Available at http://www.democrats.org/about/party_platform.

The 2008 Republican Platform. Available at http://www.gop.com/2008Platform/.

Whaples, R. and J.C. Heckman (2005). Public Choice Economics: Where is There Consensus? *American Economist* 49(1):66-78.

No author.. 2010. "A Study in Paralysis; America's Democracy. *The Economist* 384(8670):24

A SYSTEMATIC PRESENTATION OF EQUILIBRIUM BIDDING STRATEGIES TO UNDERGRADUATE STUDENTS

Felix Munoz-Garcia, Washington State University

ABSTRACT

This paper provides a non-technical introduction to auction theory. Despite the rapidly expanding literature using auction theory, and the few textbooks introducing it to upper-level Ph.D. students, the explanation in most undergraduate textbooks is very obscure and incomplete. This paper offers an introduction to auctions, analyzing optimal bidding behavior in first- and second-price auctions, and finally examines bidding strategies in common-value auctions and the winner's curse. Unlike graduate textbooks on auction theory, the paper only assumes a basic knowledge of algebra and calculus, and uses worked-out examples and figures, thus making the explanation accessible for both Economics and Business majors.

Keywords: Auction theory; First-price auction; Second-price auction; Common-value auctions; Bidding strategies.

JEL Classification: D44, D8, C7.

INTRODUCTION

Auctions have always been a large part of the economic landscape, with some auctions reported as early as in Babylon in about 500 B.C. and during the Roman Empire, in 193 A.D.[3] Auctions with precise set of rules emerged in 1595, where the Oxford English Dictionary first included the entry; and auctions houses like Sotheby's and Christie's were founded as early as 1744 and 1766, respectively. Commonly used auctions nowadays, however, are often online, with popular websites such as eBay, with US$11 billion in total revenue and more than 27,000 employees worldwide, which attracted the entry of several competitors into the online auction industry, such as QuiBids recently.

Auctions have also been used by governments throughout history. In addition to auctioning off treasury bonds, in the last decade governments started to sell air waves (3G technology). For instance, the British 3G telecom licenses generated Euro 36 billion in what British economists called "the biggest auction ever," and where several game theorists played an important role in designing and testing the auction format before its final implementation. In fact, the specific design of 3G auctions created a great controversy in most European countries during

the 1990s since, as the following figure from McKinsey (2002) shows, countries with similar population collected enormously different revenues from the sale, thus suggesting that some countries (such as Germany and the UK) better understood bidders' strategic incentives when participating in these auctions, while others essentially overlooked these issues, e.g., Netherlands or Italy.

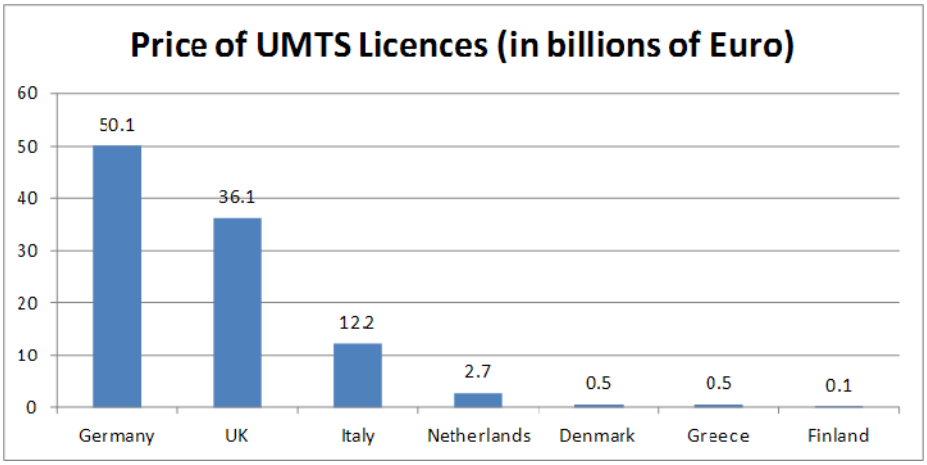


Fig 1. Prices of 3G licences.

Despite the rapidly expanding literature using auction theory, only a few graduate-level textbooks about this topic have been published; such as Krishna (2002), Milgrom (2004), Menezes and Monteiro (2004) and Klemperer (2004). These textbooks, however, introduce auction theory to upper-level (second year) Ph.D. students, using advanced mathematical statistics and, hence, are not accessible for undergraduate students. In addition, most undergraduate textbooks do not cover the topic, or present short verbal descriptions about it; see, for instance, Pindyck and Rubinfeld (2012) pp. 516-23, Perloff (2011) pp. 462-66, or Besanko and Braeutigam (2011) pp. 633-42.[4] In order to provide an attractive introduction to auction theory to undergraduate students, this paper only assumes a basic knowledge of algebra and calculus, and uses worked-out examples and figures. As a consequence, the explanations are appropriate for intermediate microeconomics and game theory courses, both for economics and business majors. In particular, the paper emphasizes the common ingredients in most auction formats (understanding them as allocation mechanism). Then, it analyzes optimal bidding behavior in first-price auctions (section three) and in second-price auctions (section four). Finally, section five examines bidding strategies in common-value auctions and the winner's curse.

AUCTIONS AS ALLOCATION MECHANISMS

Consider *N* bidders who seek to acquire a certain object, where each bidder *i* has a valuation v_i for the object, and assume that there is one seller. Note that we can design many different rules for the auction, following the same auction formats we commonly observe in real life settings. For instance, we could use:

1. First-price auction (FPA), whereby the winner is the bidder submitting the highest bid, and he/she must pay the highest bid (which in this case is his/hers).
2. Second-price auction (SPA), where the winner is the bidder submitting the highest bid, but in this case he/she must pay the second highest bid.
3. Third-price auction, where the winner is still the bidder submitting the highest bid, but now he/she must pay the third highest bid.
4. All-pay auction, where the winner is still the bidder submitting the highest bid, but in this case every single bidder must pay the price he/she submitted.

Importantly, several features are common in the above auction formats, implying that all auctions can be interpreted as allocation mechanisms with two main ingredients:

a) An *allocation rule*, specifying who gets the object. The allocation rule for most auctions determines that the object is allocated to the bidder submitting the highest bid. This was, in fact, the allocation rule for all four auction formats considered above. However, we could assign the object by using a lottery, where the probability of winning the object is a ratio of my bid relative to the sum of all bidders' bids, i.e., $prob(win)=\frac{b_i}{b_1+b_2+\cdots+b_N}$, an allocation rule often used in certain Chinese auctions.

b) A *payment rule*, which describes how much every bidder must pay. For instance, the payment rule in the FPA determines that the individual submitting the highest bid pays his own bid, while everybody else pays zero. In contrast, the payment rule in the SPA specifies that the individual submitting the highest bid (the winner) pays the second-highest bid, while everybody else pays zero. Finally, the payment rule in the all-pay auction determines that every individual must pay the bid that he/she submitted.[5]

Privately observed valuations

Before analyzing equilibrium bidding strategies in different auction formats, note that auctions are strategic settings where players must choose their strategies (i.e., how much to bid) in an incomplete information context.[6] In particular, every bidder knows his/her own valuation for the object, v_i, but does not observe other bidder *j*'s valuation, $j \neq i$. That is, bidder *i* is "in the dark" about his opponent's valuation.

Despite not observing *j*'s valuation, bidder *i* knows the probability distribution behind bidder *j*'s valuation. For instance, v_j can be relatively high, e.g., $v_j = 10$, with probability 0.4, or low, $v_j = 5$, otherwise (with probability 0.6). More generally, bidder *j*'s valuation, v_j, is distributed according to a cumulative distribution function $F(v) = prob(v_j < v)$, intuitively representing that the probability that v_j lies below a certain cutoff *v* is exactly *F*(*v*). For simplicity, we normally assume that every bidder's valuation for the object is drawn from a uniform distribution function between 0 and 1, i.e., $v_j U[0,1]$.[7] The following figure illustrates this uniform distribution where the horizontal axis depicts v_j and the vertical axis measures the cumulated probability $F(v)$. For instance, if bidder *i*'s valuation is *v*, then all points to the left-hand side of *v* in the horizontal axis represent that $v_j < v$, entailing that bidder *j*'s valuation is lower than bidder *i*'s. The mapping of these points into the vertical axis gives us the probability $prob(v_j < v) = F(v)$ which, in the case of a uniform distribution entails $F(v) = v$.[8] Similarly, the valuations to the right-hand side of *v* describe points where $v_j > v$ and, thus, bidder *j*'s valuation is higher than that of bidder *i*. Mapping these points into the vertical axis we obtain the probability $prob(v_j > v) = 1 - F(v)$ which, under a uniform distribution, implies $1 - F(v) = 1 - v$.

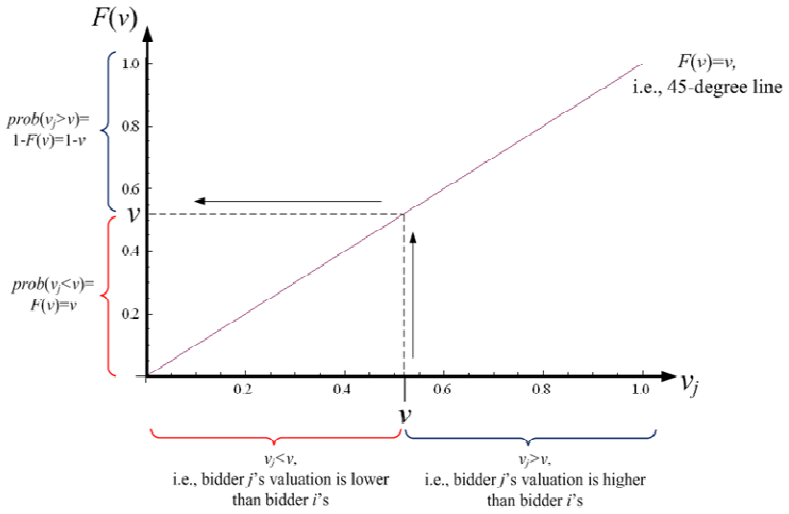


Fig 2. Uniform probability distribution.

Importantly, since all bidders are ex-ante symmetric, they will all be using the same bidding function, $b_i: [0,1] \rightarrow \mathbb{R}_+$, which maps bidder *i*'s valuation, $v_i \in [0,1]$, into a precise bid, $b_i \in \mathbb{R}_+$. However, the fact that bidders use a symmetric function does not imply that all of them submit the same bid. Indeed, depending on his privately observed valuation for the object, bidding function $b_i(v_i)$ prescribes that bidders can submit different bids. As an example, consider a symmetric bidding function $b_i(v_i) = \frac{v_i}{2}$. Hence, a bidder with valuation $v_i = 0.4$ will submit a bid of $b_i(0.4) = \frac{0.4}{2} = \$0.2$, while a different bidder whose valuation is $v_i = 0.9$ would submit a bid of $b_i(0.9) = \frac{0.9}{2} = \$0.45$. In other words, even if bidders are *symmetric* in the bidding function they use, they can be *asymmetric* in the actual bid they submit.

FIRST-PRICE AUCTIONS

Let's start analyzing equilibrium bidding behavior in the first-price auction (FPA). First, note that submitting a bid above one's valuation, $b_i > v_i$, is a dominated strategy. In particular, the bidder would obtain a negative payoff if winning, since his expected utility from participating in the auction

$$EU_i(b_i|v_i) = prob(win) \cdot (v_i - b_i) + prob(lose) \cdot 0$$

would be negative, since $v_i < b_i$, regardless of his probability of winning. Note that in the above expected utility, we specify that, upon winning, bidder *i* receives a net payoff of $v_i - x$, i.e., the difference between his true valuation for the object and the bid he submits (which ultimately constitutes the price he pays for the good if he were to win).[9] Similarly, submitting a bid b_i that exactly coincides with one's valuation, $b_i = v_i$, also constitutes a dominated strategy since, even if the bidder happens to win, his expected utility would be zero, i.e., $EU_i(b_i|v_i) = prob(win) \cdot (v_i - b_i)$, given that $b_i = v_i$. Therefore, the equilibrium bidding strategy in a FPA must imply a bid below one's valuation, $b_i < v_i$. That is, bidders must practice what is usually referred to as "bid shading." In particular, if bidder *i*'s valuation is v_i, his bid must be a share of his true valuation, i.e., $b_i(v_i) = a \cdot v_i$, where $a \in (0,1)$. The following figure illustrates bid shading in the FPA, since bidding strategies must lie below the 45-degree line.

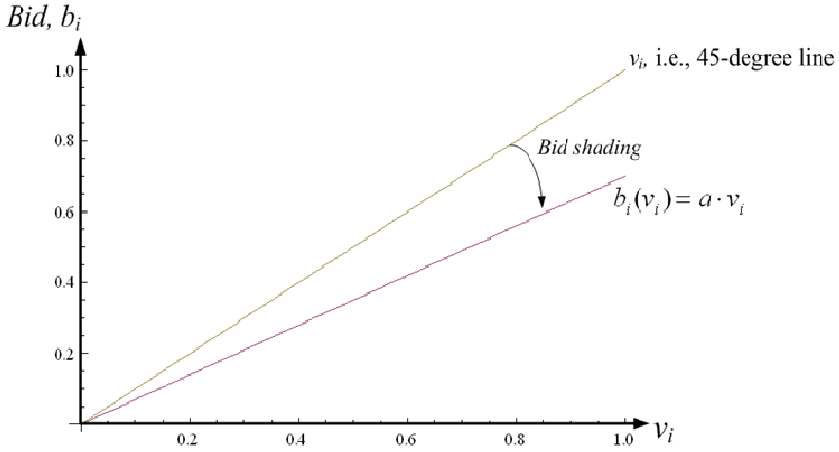


Fig 3. "Bid shading" in the FPA.

A natural question at this point is: How intense bid shading must be in the FPA? Or, alternatively, what is the precise value of the bid shading parameter *a*? In order to answer such question, we must first describe bidder *i*'s expected utility from submitting a given bid *x*, when his valuation for the object is v_i,

$$EU_i(x|v_i) = prob(win) \cdot (v_i - x) + prob(lose) \cdot 0$$

Before continuing our analysis, we still must precisely characterize the probability of winning in the above expression, i.e., $prob(win)$. Specifically, upon submitting a bid $b_i = x$, bidder *j* can anticipate that bidder *i*'s valuation is $\frac{x}{a}$, by just inverting the bidding function $b_i(v_i) = x = a \cdot v_i$, i.e., solving for v_i in $x = a \cdot v_i$ yields $v_i = \frac{x}{a}$. This inference is illustrated in the figure below where bid *x* in the vertical axis is mapped into the bidding function $a \cdot v_i$, which corresponds to a valuation of $\frac{x}{a}$ in the horizontal axis. Intuitively, for a bid *x*, bidder *j* can use the symmetric bidding function $a \cdot v_i$ to "recover" bidder *i*'s valuation, $\frac{x}{a}$.

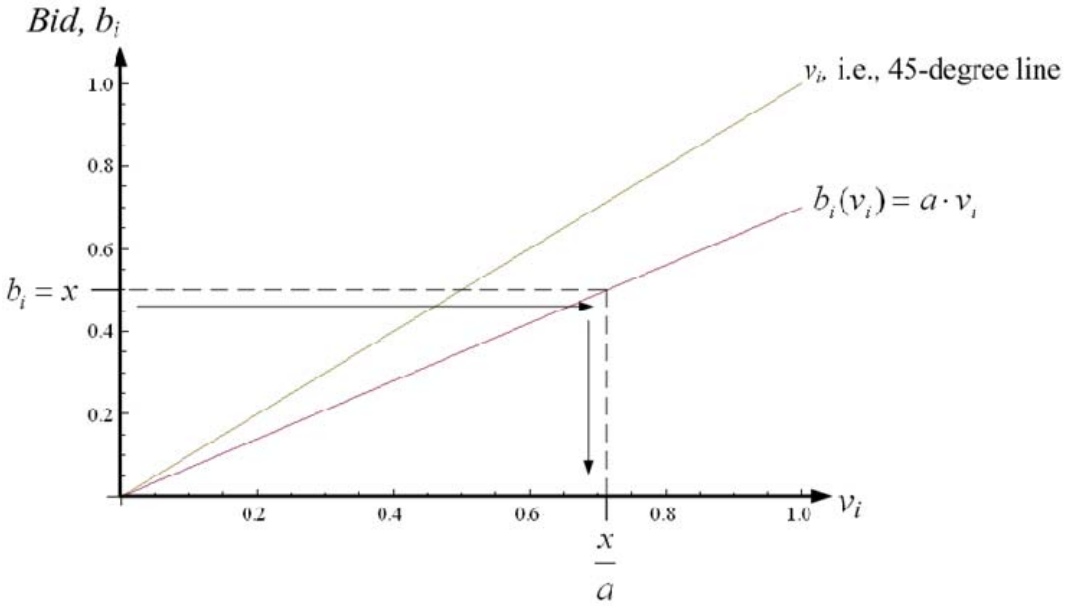


Fig 4. "Recovering" bidder *i*'s valuation.

Hence, the probability of winning is given by $prob(b_i \geq b_j)$ and, according to the vertical axis in the previous figure, $prob(b_i > b_j) = prob(x > b_j)$. If, rather than describing probability $prob(x > b_j)$ from the point of view of bids (see shaded portion of the vertical axis in figure 5 below), we characterize it from the point of view of valuations (in the shaded segment of the horizontal axis), we obtain that $prob(b_i > b_j) = prob\left(\frac{x}{a} > v_j\right)$.

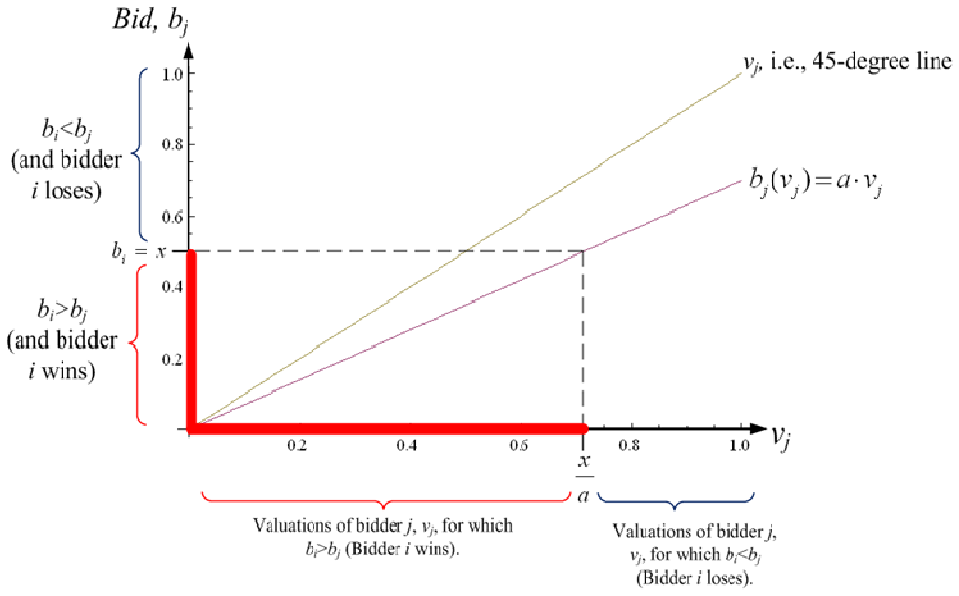


Fig. 5. Probability of winning in the FPA.

Indeed, the shaded set of valuations in the horizontal axis illustrates valuations of bidder *j*, v_j, for which his bid lies below player *i*'s bid *x*. In contrast, valuations v_j satisfying $v_j > \frac{x}{\alpha}$ entail that player *j*'s bids would exceed *x*, implying that bidder *j* wins the auction. Hence, if the probability that bidder *i* wins the object is given by $prob\left(\frac{x}{\alpha} > v_j\right)$, and valuations are uniformly distributed, then $prob\left(\frac{x}{\alpha} > v_j\right) = \frac{x}{\alpha}$.[10] We can now plug this probability of winning into bidder *i*'s expected utility from submitting a bid of x in the FPA, as follows

$$EU_i(x|v_i) = \frac{x}{\alpha}(v_i - x) = \frac{v_i x - x^2}{\alpha}$$

Taking first-order conditions[11] with respect to bidder *i*'s bid, *x*, we obtain $\frac{v_i - 2x}{\alpha} = 0$ which, solving for *x* yields bidder *i*'s optimal bidding function $x(v_i) = \frac{1}{2}v_i$. Intuitively, this bidding function informs bidder *i* how much to bid, as a function of his privately observed valuation for the object, v_i. For instance, when $v_i = 0.75$, his optimal bid is $\frac{1}{2}0.75 = 0.375$. This bidding function implies that, when competing against another bidder *j*, and only $N = 2$ players participate in the FPA, bidder *i* shades his bid in half, as the following figure illustrates.

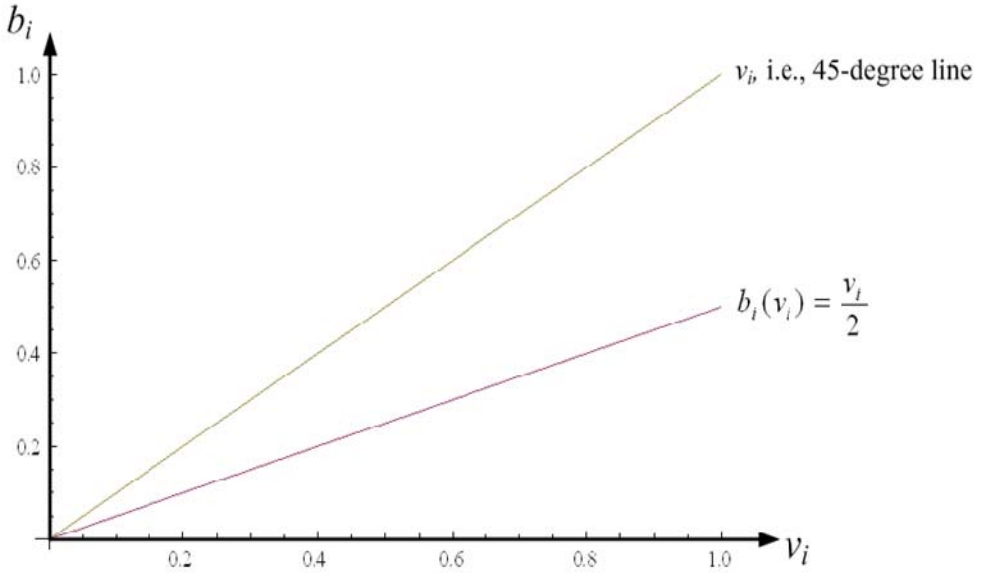


Fig 6. Optimal bidding function with *N=2* bidders.

Extending the first-price auction to *N* bidders

Our results are easily extensible to FPA with *N* bidders. In particular, the probability of bidder *i* winning when submitting a bid of $*x* is

$$prob(win) = prob\left(\frac{x}{a} > v_1\right) \cdot \ldots \cdot prob\left(\frac{x}{a} > v_{i-1}\right) \cdot prob\left(\frac{x}{a} > v_{i+1}\right) \cdot \ldots \cdot prob\left(\frac{x}{a} > v_N\right)$$
$$= \frac{x}{a} \cdot \ldots \cdot \frac{x}{a} \cdot \frac{x}{a} \cdot \ldots \cdot \frac{x}{a} = \frac{x^{N-1}}{a}$$

where we evaluate the probability that the valuation of all other *N*-1 bidders, $v^1, v^2, \ldots, v_{i-1}, v_{i+1}, \ldots, v_N$ (expect for bidder *i*) lies above the valuation $v_i = \frac{x}{a}$ that generates a bid of exactly *x* dollars. Hence, bidder *i*'s expected utility from submitting *x* becomes

$$EU_i\ (x|v_i) = \left(\frac{x}{a}\right)^{N-1} (v_i - x) + \left[1 - \left(\frac{x}{a}\right)^{N-1}\right] 0$$

Taking first-order conditions with respect to his bid, *x*, we obtain

$$-\left(\frac{x}{a}\right)^{N-1} + \left(\frac{x}{a}\right)^{N-2} \left(\frac{1}{a}\right) (v_i - x) = 0$$

Rearranging, $\left(\frac{x}{a}\right)^N \frac{a}{x^2}[(N-1) - nx] = 0$, and solving for *x*, we find bidder *i*'s optimal bidding function, $x(v_i) = \frac{N-1}{N} v_i$. The following figure depicts the bidding function for the case of *N*=2, *N*=3, and *N*=4 bidders, showing that bid shading is ameliorated when more bidders participate in the auction, i.e., bidding functions approach the 45-degree line. Indeed, for *N*=2 the optimal bidding function is $\frac{1}{2} v_i$, but it increases to $\frac{2}{3} v_i$ when *N*=3 bidders compete for the object, to $\frac{3}{4} v_i$ when *N*=4 players participate in the auction, etc. For an extremely large number of bidders, e.g., *N*=2,000, bidder *i*'s optimal bidding function becomes $b_i(v_i) = \frac{1,999}{2,000} v_i \approx v_i$ and, hence, bidder *i*'s bid almost coincides with his valuation for the object, describing a bidding function that approaches the 45-degree line.

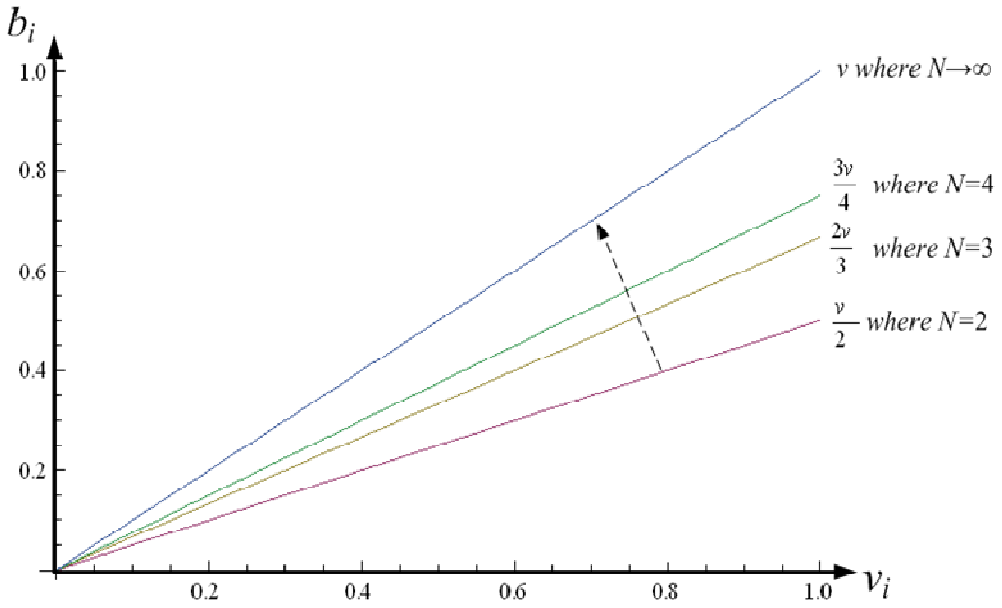


Fig 7. Optimal bidding function increases in *N*.

Intuitively, if bidder *i* seeks to win the object, he can shade his bid when only another bidder competes for the good, since the probability of him assigning a large valuation to the object is relatively low. However, when several players compete in the auction, the probability that some of them have a high valuation for the object (and, thus submits a high bid) increases. That is, competition gets "tougher" as more bidders participate and, as a consequence, every bidder must increase his bid, ultimately ameliorating his incentives to practice bid shading.

First-price auctions with risk-averse bidders

Let us next analyze how our equilibrium results would be affected if bidders are risk averse, i.e., their utility function is concave in income, *x*, e.g., $u(x) = x^{\alpha}$, where $0 < \alpha \leq 1$ denotes bidder *i*'s risk-aversion parameter. In particular, when $\alpha = 1$ he is risk neutral, while when α decreases, he becomes risk averse.[12] First, note that the probability of winning is unaffected, since, for a symmetric bidding function $b_i(v_i) = \alpha \cdot v_i$ for every bidder *i*, where $\alpha \in (0,1)$, the probability that bidder *i* wins the auction against another bidder *j* is

$$prob(b_i \geq b_j) = prob(x \geq b_j) = prob\left(\frac{x}{\alpha} \geq v_j\right) = \frac{x}{\alpha}$$

Therefore, bidder *i*'s expected utility from participating in this auction is

$$EU_i(x|v_i) = \frac{x}{a} \times (v_i - x)^{\alpha} + \left(1 - \frac{x}{a}\right) \times 0$$

where, relative to the case of risk-neutral bidders analyzed above, the only difference arises in the evaluation of the net payoff from winning, $v_i - x$, which it is evaluated as $(v_i - x)^{\alpha}$. Taking first-order conditions with respect to his bid, *x*, we have

$$\frac{1}{a}(v_i - x)^{\alpha} - \frac{x}{a}\alpha(v_i - x)^{\alpha-1} = 0,$$

and solving for *x*, we find the optimal bidding function, $x(v_i) = \frac{v_i}{1+\alpha}$. Importantly, this case embodies that of risk-neutral bidders analyzed above as a special case. Specifically, when $\alpha = 1$, bidder *i*'s optimal bidding function becomes $x(v_i) = \frac{v_i}{2}$. However, when his risk aversion increases, i.e., α decreases, bidder *i*'s optimal bidding function increases. Specifically,

$$\frac{\partial x(v_i)}{\partial \alpha} = -\frac{v_i}{(1 \quad \alpha)^2}$$

which is negative for all parameter values. In the extreme case in which α decreases to $\alpha \to 0$, the optimal bidding function becomes $x(v_i) = v_i$, and players do not practice bid shading. The following figure illustrates the increasing pattern in players' bidding function, starting from $\frac{v_i}{2}$ when bidders are risk neutral, $\alpha = 1$, and approaching the 45-degree line (no bid shading) as players become more risk averse.

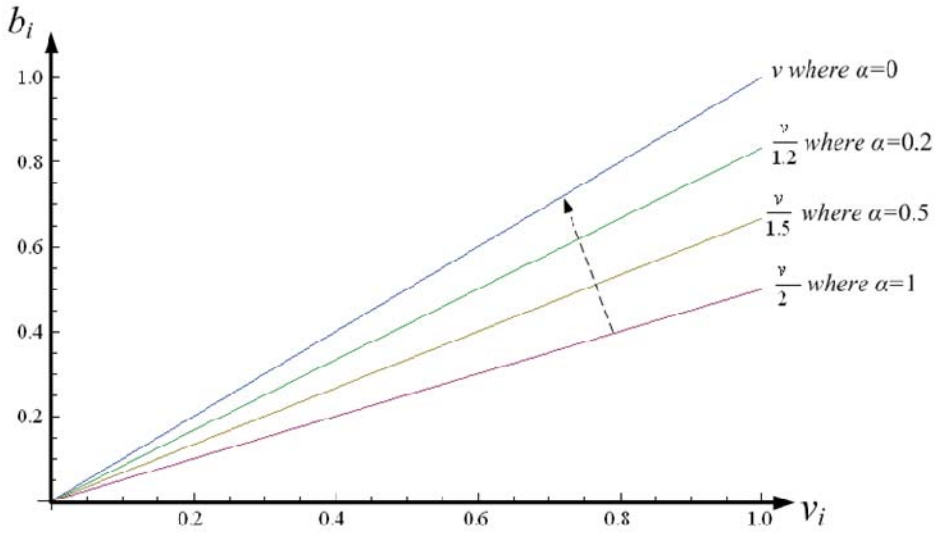


Fig. 8. Optimal bidding function with risk-averse bidders.

Intuitively, a risk-averse bidder submits more aggressive bids than a risk-neutral bidder in order to minimize the probability of losing the auction. In particular, consider that bidder *i* reduces his bid from b_i to $b_i - \varepsilon$. In this context, if he wins the auction, he obtains an additional profit of ε, since he has to pay a lower price for the object he acquires. However, by lowering his bid, he increases the probability of losing the auction. Importantly, for a risk-averse bidder, the positive effect of slightly lowering his bid, arising from getting the object at a cheaper price, is offset by the negative effect of increasing the probability that he loses the auction. In other words, since the possible loss from losing the auction dominates the benefit from acquiring the object at a cheaper price, the risk-averse bidder does not have incentives to reduce his bid, but rather to increase it, relative to the risk-neutral bidders.

SECOND-PRICE AUCTION

In this class of auctions, bidding your own valuation, i.e., $b_i(v_i) = v_i$, is a weakly dominant strategy for all players. That is, regardless of the valuation you assign to the object, and independently on your opponents' valuations, submitting a bid $b_i(v_i) = v_i$ yields expected profit equal or above that from submitting any other bid, $b_i(v_i) \neq v_i$. In order to show this bidding strategy is an equilibrium outcome of the SPA, let's first examine bidder *i*'s expected payoff from submitting a bid that coincides with his own valuation v_i (which we refer to as the First case below), and then compare it with what he would obtain from deviating to bids below his valuation for the object, $b_i(v_i) < v_i$ (denoted as Second case), or above his valuation, $b_i(v_i) > v_i$ (Third case). Let us next separately analyze the payoffs resulting from each bidding strategy.

First case: If the bidder submits his own valuation, $b_i(v_i) = v_i$, then either of the following situations can arise (for presentation purposes, the figure below depicts each of the three cases separately):

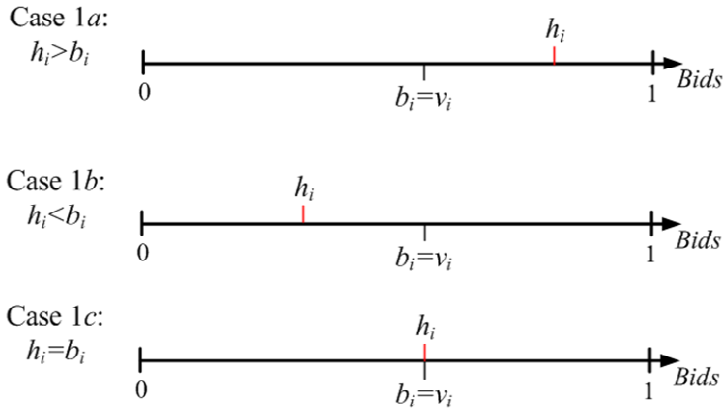


Fig 9. Cases arising when $b_i(v_i) = v_i$.

a) If his bid lies below the highest competing bid, i.e., $b_i < h_i$ where $h_i = \max_{j \neq i}\{b_j\}$,[13] then bidder *i* loses the auction, obtaining a zero payoff.

b) If his bid lies above the highest competing bid, i.e., $b_i > h_i$, then bidder *i* wins the auction. In this case, he obtains a net payoff of $v_i - h_i$, since in a SPA the winning bidder does not have to pay the bid he submitted, but the second-highest bid, which is h_i in this case since $b_i > h_i$.

c) If, instead, his bid coincides with the highest competing bid, i.e., $b_i = h_i$, then a tie occurs. For simplicity, ties are normally solved in auctions by randomly assigning the object to the bidders who submitted the highest bids. As a consequence, bidder *i*'s payoff becomes $v_i - h_i$, but with only $\frac{1}{2}$ probability, i.e., his expected payoff $\frac{1}{2}(v_i - h_i)$.[14]

Second case: Let us now compare the above equilibrium payoffs with those bidder *i* could obtain by deviating towards a bid that shades his valuation, i.e., $b_i < v_i$. In this case, we can also identify three cases emerging (see figure 10), depending on the ranking between bidder *i*'s bid, b_i, and the highest competing bid, h_i.

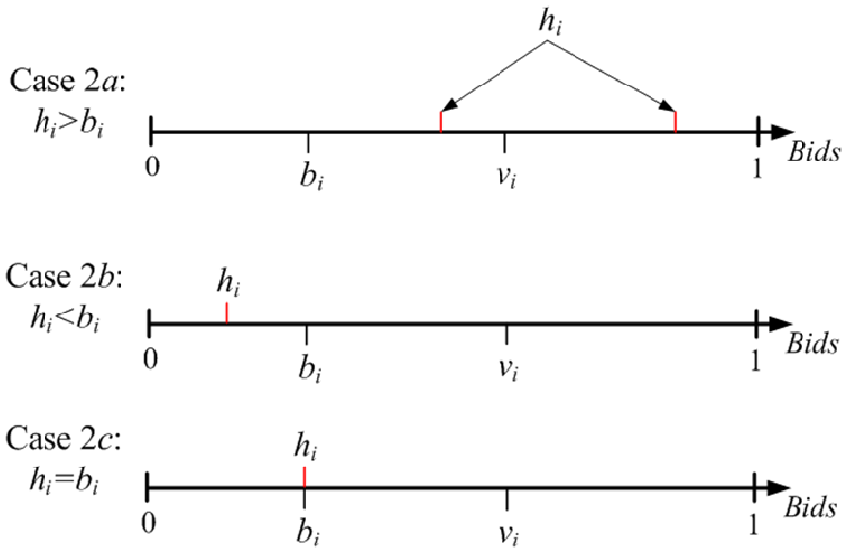


Fig 10. Cases arising when $b_i(v_i) < v_i$.

a) If his bid lies below the highest competing bid, i.e., $b_i < h_i$, then bidder *i* loses the auction, obtaining a zero payoff.

b) If his bid lies above the highest competing bid, i.e., $b_i > h_i$, then bidder *i* wins the auction, obtaining a net payoff of $v_i - h_i$.

c) If, instead, his bid coincides with the highest competing bid, i.e., $h_i = h_j$, then a tie occurs, and the object is randomly assigned, yielding an expected payoff of $\frac{1}{2}(v_i - h_i)$.

Hence, we just showed that bidder *i* obtains the same payoff submitting a bid that coincides with his privately observed valuation for the object $b_i = v_i$, as in the First case) and shading his bid $b_i < v_i$, as described in teh Second case). Therefore, he does not have incentives to conceal his bid, since his payoff would not improve from doing so.

Third case: Let us finally examine bidder *i*'s equilibrium payoff from submitting a bid above his valuation, i.e., $b_i(v_i) > v_i$. In this case, three cases also arise (see figure 11).

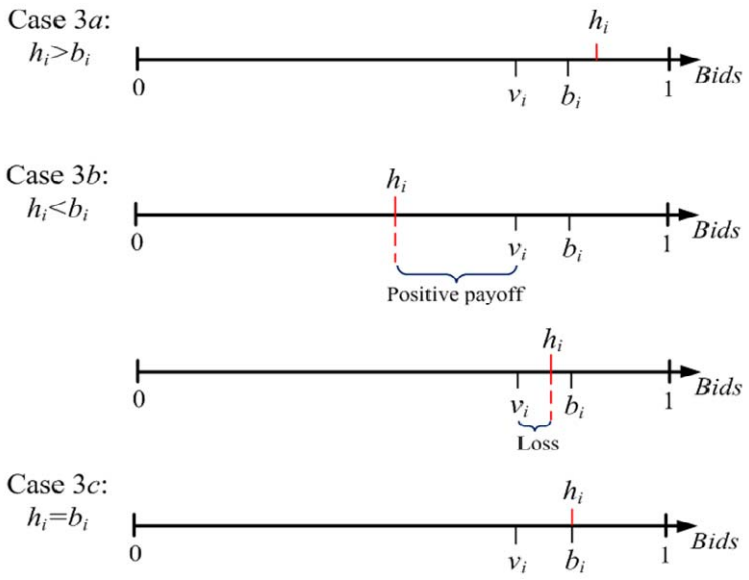


Fig 11. Cases arising when $b_i(v_i) > v_i$.

a) If his bid lies below the highest competing bid, i.e., $b_i < h_i$, then bidder *i* loses the auction, obtaining a zero payoff.

b) If his bid lies above the highest competing bid, i.e., $b_i \geq h_i$, then bidder i wins the auction. In this scenario, his payoff becomes $v_i - h_i$, which is positive if $v_i \geq h_i$, or negative otherwise. (These two situations are depicted in case 3b of figure 11.) The latter case, in which bidder *i* wins the auction but at a loss (negative expected payoff), did not exist in our above analysis of $b_i(v_i) = v_i$ and $b_i(v_i) < v_i$, since players did not submit bids above their own valuation. Intuitively, the possibility of a negative payoff arises because bidder *i*'s valuation can lie below the second-highest bid, as figure 11 illustrates, where $v_i < h_i < b_i$.

c) If, instead, his bid coincides with the highest competing bid, i.e., $b_i = h_i$, then a tie occurs, and the object is randomly assigned, yielding an expected payoff of $\frac{1}{2}(v_i - h_i)$. Similarly as our above discussion, this expected payoff is positive if $v_i \geq h_i$, but negative otherwise.

Hence, bidder *i*'s payoff from submitting a bid above his valuation either coincides with his payoff from submitting his own value for the object, or becomes strictly lower, thus nullifying his incentives to deviate from his equilibrium bid of $b_i(v_i) = v_i$. In other words, there is no bidding strategy that provides a strictly higher payoff than $b_i(v_i) = v_i$ in the SPA, and all players bid their own valuation, without shading their bids; a result that differs from the optimal bidding function in FPA, where players shade their bids unless $N \to \infty$.

Remark. The above equilibrium bidding strategy in the SPA is, importantly, unaffected by the number of bidders who participate in the auction, *N*, or their risk-aversion preferences. In particular, our above discussion considered the presence of *N* bidders, and an increase in their number does not emphasize or ameliorate the incentives that every bidder has to submit a bid that coincides with his own valuation for the object, $b_i(v_i) = v_i$. Furthermore, the above results remain when bidders evaluate their net payoff, e.g., $v_i - h_i$, according to a concave utility function, such as $u(x) = x^{\alpha}$, exhibiting risk aversion. Specifically, for a given value of the highest competing bid, h_i, bidder *i*'s expected payoff from submitting a bid $b_i(v_i) = v_i$ would still be weakly larger than from deviating to a bidding strategy above, $b_i(v_i) > v_i$, or below, $b_i(v_i) < v_i$, his true valuation for the object.

Efficiency in auctions

Auctions, and generally allocation mechanism, are characterized as efficient if the bidder (or agent) with the highest valuation for the object is indeed the person receiving the object. Intuitively, if this property does not hold, the outcome of the auction (i.e., the allocation of the object) would open the door to negotiations and arbitrage among the winning bidder —who, despite obtaining the object, is not the player who assigns the highest value to it— and other bidder/s with higher valuations who would like to buy the object from him. In other words, the auction's outcome would still allow for negotiations that are beneficial for all parties involved, i.e., usually referred as Pareto improving negotiations, thus suggesting that the initial allocation was not efficient.

According to this criterion, both the FPA and the SPA are efficient, since the bidder with the highest valuation submits the highest bid, and the object is ultimately assigned to the player who submits the highest bid. Other auction formats, such as the Chinese (or lottery) auction described in the Introduction, are not necessarily efficient, since they may assign the object to an individual who did not submit the highest valuation for the object. In particular, recall that the probability of winning the object in this auction is a ratio of the bid you submit relative to the sum of all players' bids. Hence, a bidder with a low valuation for the object, and who submits the lowest bid, e.g., $1, can still win the auction. Alternatively, the person that assigns the highest value to the object, despite submitting the highest bid, might not end up receiving the object for sale. Therefore, for an auction to satisfy efficiency, bids must be increasing in a player's valuation, and the probability of winning the auction must be one (100%) if a bidder submits the highest bid.

COMMON-VALUE AUCTIONS

The auction formats considered above assume that each bidders privately observes his own valuation for the object, and this valuation is distributed according to a distribution function $F(v)$, e.g., a uniform distribution, implying that two bidders are unlikely to assign the same value to the object for sale. However, in some auctions, such as the government sale of oil leases, bidders (oil companies) might assign the same monetary value to the object (common value), i.e., the profits they would obtain from exploiting the oil reservoir. Bidders are, nonetheless, unable to precisely observe the value of this oil reservoir but, instead, gather estimates of its value. In the oil lease example, firms cannot accurately observe the exact volume of oil in the reservoir, or how difficult it will be to extract, but can accumulate different estimates from their own engineers, or from other consulting companies, that inform the firm about the potential profits to be made from the oil lease. Such estimates are, nonetheless, imprecise, and only allow the firm to assign a value to the object (profits from the oil lease) within a relatively narrow range, e.g., $v \in [10,11,\ldots,20]$ in millions of dollars. Consider that oil company A hires a consultant, and gets a signal (a report), s, as follows

$$s - \begin{cases} v + 1 \text{ with prob.} \frac{1}{2}, \text{and} \\ v - 2 \text{ with prob.} \frac{1}{2} \end{cases}$$

and, hence, the signal is above the true value to the oil lease with 50% probability, or below its value otherwise. We can alternatively represent this information by examining the conditional probability that the true value of the oil lease is *v*, given that the firm receives a signal *s*, is

$$prob(v|s) - \begin{cases} \frac{1}{2} \text{ if } v = s - 2 \text{ (overestimate), and} \\ \frac{1}{2} \text{ if } v = s + 2 \text{ (underestimate)} \end{cases}$$

since the true value of the lease is overestimated when $v = s - 2, i.e., s = v + 2$ and the signal is above *v*; and underestimated when $v = s + 2, i.e., s = v - 2$ and the signal lies below *v*. Notice that, if company *A* was not participating in the auction, then the expected value of the oil lease would be

$$\frac{1}{2}(s-2) + \frac{1}{2}(s+2) = \frac{(s-2)+(s+2)}{2} = s$$

implying that the firm would pay for the oil lease a price $p \ll s$, making a positive expected profit. But, what if the oil company participates in a FPA for the oil lease against another company *B*? In this context, every firm uses a different consultant, i.e., can receive different signals, but does not know whether their consultant systematically over- or under-estimates the true value of the oil lease. In particular, consider that every firm receives a signal *s* from their consultant. Observing its own signal, but not observing the signal received by the other firm, every firm $i=\{A,B\}$ submits a bid from the set {1,2,…,20}, where the upper bound of this interval represents the maximum value of the oil lease according to all estimates.

We will next show that slightly shading your bid, e.g., submitting $b = s - 1$, cannot be optimal for any firm. At first glance, however, such a bidding strategy seems sensitive: the firm bid is increasing in the signal it receives and, in addition, its bid is below the signal, $b \ll s$, entailing that, if the true value of the oil lease was s, the firm would obtain a positive expected profit from winning. In order to show that bid $b = s - 1$ cannot be optimal, consider that firm *A* receives a signal $s = 10$, and thus submits a bid $b = s - 1 = 10 - 1 = \$9$. Given such a signal, the true value of the oil lease is

$$v - \begin{cases} s + 2 = 12 \ \textit{with prob.} \dfrac{1}{2}, \textit{and} \\ s - 2 = 8 \ \textit{with prob.} \dfrac{1}{2} \end{cases}$$

Specifically, when the true value of the oil lease is v=12, firm A receives a signal of $s_A = 10$ (an underestimation of the true valuation, 12), while firm B receives a signal of $s_B = 14$ (an overestimation). In this setting, firms bid $b_A = 10 - 1 = \$9$ and $b_B = 14 - 1 = \$13$ and, thus, firm A loses the auction. If, in contrast, the true value of the lease is v =8, firm A receives a signal of $s_A = 10$ (an overestimation of the true valuation, 8), while firm B receives a signal $s_B = 6$ (an underestimation). In this context, firms bid $b_A = 10 - 1 = \$9$, and $b_B = 6 - 1 = \$5$, and firm A wins the auction. In particular, firm A's expected profit from participating in this auction is

$$\frac{1}{2}(8 - 9) + \frac{1}{2}0 = -\frac{1}{2}$$

which is negative! This is the so-called "winner's curse" in common-value auctions. In particular, the fact that a bidder wins the auction just means that he probably received an overestimated signal of the true value of the object for sale, as firm A receiving signal $s_A = 10$ in the above example. Therefore, in order to avoid the winner's curse, participants in common-value auctions must significantly shade their bid, e.g., b=s-2 or less, in order to consider the possibility that the signals they receive are overestimating the true value of the object.[15]

The winner's curse in practice. Despite the straightforward intuition behind this result, the winner's curse has been empirically observed in several controlled experiments. A common example is that of subjects in an experimental lab, where they are asked to submit bids in a common-value auction where a jar of nickels is being sold. Consider that the instructor of one of your courses comes to class with a jar plenty of nickels. The monetary value you assign to the jar coincides with that of your classmates, i.e., its value is common, but none of you can accurately estimate the number of nickels in the jar, since you can only gather some imprecise information about its true value by looking at the jar for a few seconds. In these experiments, it is usual to find that the winner ends up submitting a bid a monetary amount beyond the jar's true value, i.e., the winner's curse emerges. (For some experimental evidence on the winner's curse see, for instance, Thaler (1988).)

More surprisingly, the winner's curse has also been shown to arise among oil company executives. Hendricks et al. (2003) analyze the bidding strategies of companies, such as Texaco, Exxon, an British Petroleum, when competing for the mineral rights to properties 3-200 miles

off-shore and initially owned by the U.S. government. Generally, executives did not systematically fall prey of the winner's curse, since their bids were about one third of the true value of the oil lease. As a consequence, if their bids resulted in their company winning the auction, their expected profits would become positive. Texaco executives, however, not only fell prey of the winner's curse, but submitted bids above the estimated value of the oil lease. Such a high bid, if winning, would have resulted in negative expected profits. One cannot help but wonder if Texaco executives were enrolled in a remedial course on auction theory.

ENDNOTES

1. I thank Ana Espinola-Arredondo for her insightful suggestions, and Jesse Fosse and Donald Petersen for their helpful assistance.
2. 103G Hulbert Hall, School of Economic Sciences, Washington State University, Pullman, WA, 99164-6210. Tel. 509-335-8402. Fax 509-335-1173. E-mail: fmunoz@wsu.edu.
3. In particular, the Praetorian Guard, after killing Pertinax, the emperor, announced that the highest bidder could claim the Empire. Didius Julianus was the winner, becoming the emperor for two short months, after which he was beheaded.
4. Varian's (2010) textbook provides a more complete introduction to auctions and mechanism design but, unlike this paper, it does not focus on equilibrium bidding strategies.
5. This auction format is used by the internet seller QuiBids.com. For instance, if you participate in the sale of a new iPad, and you submit a low bid of $25 but some other bidder wins by submitting a higher bid, you will still see your $25 bid withdrawn from your QuiBids account.
6. Auctions are, hence, regarded as an example of Bayesian game.
7. Note that this assumption does not imply that bidder *j* does not assign a valuation v_j larger than one to the object but, instead, that his range of valuations, e.g., from 0 to $\bar{v}$, can be normalized to the interval [0,1].
8. For more references about probability distributions and its properties, see textbooks on Statistics for Economists, such as Anderson et al (2009), McClave et al (2010), and Keller (2011). For a more rigorous treatment, see Mittlehammer (1996).
9. Upon loosing, bidders do not obtain any object and, in this auction, do not have to pay any monetary amount, thus implying a zero payoff.
10. Recall that, if a given random variable *y* is distributed according to a uniform distribution function $U[0,1]$, the probability that the value of *y* lies below a certain cutoff *c*, is exactly *c*, i.e., $prob(y < c) = F(c) = c$.
11. For standard references on calculus applied to Economics and Business, see Klein (2001), and Wainwright and Chiang (2004).
12. An example you have probably encountered in intermediate microeconomics courses includes $u(x) = \sqrt{x}$ since $\sqrt{x} = x^{1/2}$. As a practice, note that the Arrow-Pratt coefficient of absolute risk aversion $r_A(x) = -\frac{u''(x)}{u'(x)}$ for this utility function yields $\frac{1-\alpha}{x}$, confirming that, when $\alpha = 1$, the coefficient of risk aversion becomes zero, but when $0 < \alpha < 1$, the coefficient is positive.
13. Intuitively, expression $b_{-i} = \max_{j \neq i}\{b_j\}$ just finds the highest bid among all bidders different from bidder $i, j \neq i$.

14. Note that, more generally, if $K \geq 2$ bidders coincide in submitting the highest bid, their expected payoff becomes $\frac{1}{K}(v_i - b_i)$.

15. It can be formally shown that, in the case of N=2 bidders, the optimal bidding function is $b_i(v_i) = \frac{1}{2}s_i$, where s_i denotes the signal that bidder i receives. More generally, for N bidders, bidder i's optimal bid becomes $b_i(v_i) = \frac{(N-1)(N-1)}{2N^2}s_i$. For more details, see Harrington (2009), pp. 321-23.

REFERENCES

Anderson, D. R., D. J. Sweeney, and T. A. Williams (2009), *Statistics for Business and Economics*, South-Western College Publishing.

Besanko, D. and R. Braeutigam (2011) *Microeconomics*, Wiley Publishers.

Harrington, J. (2009) *Games, Strategies, and Decision Making*, Worth Publishers.

Hendricks, K., J. Pinske, and R.H. Porter (2003) "Empirical implications of equilibrium bidding in first-price, symmetric, common-value auctions," *Review of Economic Studies*, 70, pp. 115-45.

Keller, G. (2011), *Statistics for Management and Economics*, South-Western College Publishing.

Klein, M. W. (2001), *Mathematical Methods for Economics*, Addison Wesley.

Klemperer, P. (2004) *Auctions: Theory and Practice* (The Toulouse Lectures in Economics). Princeton University Press.

Krishna, V. (2002), *Auction Theory*. Academic Press.

McClave, J. T., P. G. Benson, and T. Sincich (2010), *Statistics for Business and Economics*, Pearson Publishers.

McKinsey (2002), "Comparative Assessment of the Licensing Regimes for 3G Mobile Communications in the European Union and their impact on the Mobile Communications Sector," available at the European Commission's website: http://ec.europa.eu/information_society/topics/telecoms/radiospec.

Menezes, F.M. and P.K. Monteiro (2004), *An Introduction to Auction Theory*, Oxford University Press.

Mittlehammer, R. (1996), *Mathematical Statistics for Economics and Business*, Springer.

Milgrom, P. (2004), *Putting Auction Theory to Work*, Cambridge University Press.

Pindyck, R. and D. Rubinfeld (2012) *Microeconomics*, Pearson Publishers.

Perloff, J.M. (2011) *Microeconomics, Theory and Applications with Calculus*, Addison Wesley.

Thaler, R. (1988) "Anomalies: The Winner's Curse," *The Journal of Economic Perspectives*, 2(1), pp. 191-202.

Varian, H. (2010) *Intermediate Microeconomics*, Norton Publishing.

Wainwright, K. and A. Chiang (2004), *Fundamental Methods of Mathematical Economics*, McGraw-Hill.

THE EFFECT OF JOB CHARACTERISTICS ON JOB SATISFACTION IN THE UNITED STATES AND CHINA

Kevin D. Neuman, University of Wisconsin – Stevens Point

ABSTRACT

This study utilizes a detailed data set from a regional manufacturing firm to investigate how worker satisfaction differs between the United States and China. While there does not appear to be an overall difference in satisfaction across countries, the similar levels of satisfaction appear to be driven by different factors. Chinese workers respond more positively to good communication and the communal aspect of work such as morale and work friendships. However, Chinese workers react negatively to training and fear negative management reaction to surveys.

INTRODUCTION

As companies expand their operations internationally, understanding the attitudes of their workforce becomes an increasingly difficult proposition. While difficult to attain, this understanding is important. If international workers differ from their United States counterparts in terms of their attitudes towards the job, management styles will need to adapt to the context of specific countries or productivity may suffer. There are many aspects of worker attitudes that are important, but one broad, very useful characteristic that deserves particular attention is the overall job satisfaction of workers. Overall job satisfaction has been shown to lead to better productivity (Munyon et al., 2010), fewer quits (Bockerman and Ilmakunnas, 2009; Green, 2010), less absenteeism (Drago and Wooden, 1992), and less harmful behavior on the job (Mangione and Quinn, 1975). While the general literature on job satisfaction in western economies is well developed and has a long history, one element that has been less examined is the determinants of job satisfaction in other countries, particularly the People's Republic of China. Given China's rapidly expanding economy and its role as a global manufacturer, even relatively small companies are finding opportunities in the Chinese market. Understanding the satisfaction of the Chinese worker in particular will become an essential part of managing a global organization.

This project examines the effect of job characteristics on job satisfaction for workers in the United States and China. The study contributes to the gap in the satisfaction literature in a few important ways. First, prior studies have not typically had detailed information about employee attitudes towards management styles or organizational characteristics such as communication style, training, networking opportunities, and the quality of teamwork. This

detailed information about how people feel about specific aspects of their positions and the organization will provide practical evidence about what workers truly value on the job and how they react to their managers. Additionally, although prior studies have examined some aspects of the Chinese worker, they typically do not test for differences across countries. This study explicitly compares worker responses and the determinants of those responses, to their counterparts in the United States. By examining whether there are significant cross-country differences in how workers perceive their jobs, the findings provide valuable information to inform companies about cross cultural differences between the United States and China that impact operations.

To conduct the analysis I utilize a unique dataset from a regional manufacturing firm with operations across the United States and in China. The detailed dataset contains information about basic job satisfaction, but also inquires about worker attitudes related to aspects such as teamwork, desired mobility, training, wages, and communication, to name a few. The range of questions included in the study allows me to examine a different set of attitudes than other studies. In addition, since the workers are part of the same organization I am able to control for a major variant across countries in other studies, namely that the workers examined work for different employers with different policies and management styles. In this situation, since the workers are employed by the same company, the primary difference across countries is the workers themselves.

JOB SATISFACTION LITERATURE

The prior literature on job satisfaction is extensive and cross-disciplinary (for an excellent review of both the findings and the methodologies see Linz and Semykina, 2012). Perhaps not surprisingly, one common finding is that higher absolute pay increases satisfaction (Artz, 2008; Brown et al., 2008; Clark et al., 2009; Heywood et al., 2002; Heywood and Wei, 2006; Kosteas, 2011; Pouliakas and Theodossiou, 2010; Sousa-Poza and Sousa-Poza, 2000). Higher relative pay tends to increase satisfaction as well (Brown et al., 2008; Gao and Smyth, 2009; Kosteas, 2011), although the finding is somewhat ambiguous as at least one study finds that lower than average pay actually increases satisfaction by signaling the potential for greater future wages (Clark et al., 2009). In terms of how workers are paid, incentive pay schemes tend to increase satisfaction (Artz, 2008; Heywood and Wei, 2006). Job characteristics matter as well, with workers reporting greater satisfaction with less dangerous/hazardous work (Bockerman and Ilmakunnas, 2009; Sousa-Poza and Sousa-Poza, 2000), greater opportunities for advancement (Bockerman and Ilmakunnas, 2009; Kosteas, 2011; Sousa-Poza and Sousa-Poza, 2000) and training (Artz, 2008; Gazioglu and Tansel, 2006), and in public firms (Artz, 2008). With regard to worker characteristics, non-union workers consistently report greater satisfaction (Artz, 2008; Heywood et al., 2002; Kosteas, 2011), as do young and older workers (Artz, 2008; Clark et al., 2009; Heywood et al., 2002; Linz and Semykina, 2012; Pouliakas and Theodossiou, 2010), women (Artz, 2008; Heywood et al., 2002; Kosteas, 2011), those with less education (Artz, 2008;

Heywood et al., 2002; Pouliakas and Theodossiou, 2010; Sousa-Poza and Sousa-Poza, 2000), and with shorter tenure (Heywood and Wei, 2006; Kosteas, 2011; Pouliakas and Theodossiou, 2010).

The literature on effects for Chinese workers is less extensive as traditionally most studies of job satisfaction have focused on western countries and economies. However, recently some studies have started to look at the job satisfaction question in China. As in the previous, primarily Western studies, Chinese workers have been found to be positively influenced by greater absolute income (Gao and Smyth, 2009; Nielsen and Smyth, 2008; Wang et al., 2013), greater relative income (Gao and Smyth, 2009; Leung et al., 2001), more training and promotion opportunities (Ma and Trigo, 2008; Wang et al., 2013), less dangerous and dirty work (Donald and Siu, 2001; Wang et al., 2013), and fewer work hours (Gao and Smyth, 2009; Wang et al., 2013). For demographic characteristics, older and less educated workers are also more satisfied (Nielsen and Smyth, 2008). In contrast, being a union member has a positive effect for workers in China (Gao and Smyth, 2009), although the meaning of union membership is very different across economies. Additionally, greater organizational commitment of the workers has been shown to have a positive effect (Siu, 2002; Wong et al., 2001), as does greater perceived organizational support for the worker (Rutherford et al., 2012). Generational issues have been discovered as well. The so called new generation of migrants has been found to have greater satisfaction than the traditional generation primarily due to having better working conditions (Wang et al., 2013). Overall these studies have begun to paint a picture of what determines the satisfaction of the Chinese worker.

The current study fits into this literature in a few ways. Most importantly, prior studies have not looked at as detailed of job characteristics and management behaviors as the current study. Other studies may look at the industry or occupation, or at general questions about job stressors such as work hours and work conditions, but they do not have information about the more intricate details of the job such as whether workers feel respected by management or whether enough information is communicated to them in order to properly do their job. The current study can answer these questions as well as questions about the effects of the adequacy of training, feelings of being able to contribute to the company, and of perceptions of teamwork and morale. These job characteristics and employee attitudes will help shed light on the effects of the interactions between managers and workers. Unfortunately the level of job detail does not come without a cost, as the current study does not have as detailed a set of demographic characteristics due to company concerns about confidentiality of respondents. Nevertheless the study does explore a unique set of characteristics and should provide valuable information to management in organizations with Chinese operations.

A second important contribution of the current study is that the prior studies that do examine job satisfaction in a Chinese context do not explicitly test for differences between China and Western economies. While it is reasonably easy to see differences without testing if there are significant coefficients with different signs in each country, statistical testing is important to

identify more subtle differences. For example, studies have found that greater relative wages increase satisfaction in both the United States and China, but without statistically testing it is impossible to say whether the positive effect is greater in one country than another. While both effects are statistically significant, it may be that the relative income issue is of great importance to workers in the U.S., but of relatively low practical importance to workers in China. Having knowledge of these differences across countries would be important information to have as managers may be able to tailor policies and procedures to different countries rather than applying a one size fits all policy to the organization.

There are a few cautions related to the current study that should be noted. The first is that the idea of job satisfaction may differ across countries. A few studies have found some evidence to this effect (Han et al., 2009; Kristensen and Johansson, 2008). This may mean that differences between workers across countries are not really due to workers placing different values on specific job characteristics when determining satisfaction, but rather are due to the fact that they simply perceive and understand the overall concept of job satisfaction differently. While this is a possibility and makes interpretation of the results a bit more difficult, the study still provides information about the unique set of job characteristics within each country. A second issue relates to the broader applicability of the results. As with any study in a single company, institutional rules and policies may be distinct enough from other organizations to make the extension of results difficult to other organizations and industries. While this is true, the benefit to looking at a single organization is that the institutional rules and policies will be the same for workers in both countries eliminating or at least reducing an unobservable influence that could be driving results in studies using different organizations. Once again, while the issue may complicate interpretation of the findings it does not eliminate the value of the contributions of the study.

DATA AND METHODOLOGY

The data for the study is taken from a survey administered during the fall of 2011 at a regional manufacturing firm. The firm has roughly 1,800 union and non-union production and trades workers, and 600 non-union office workers in various occupations. Although a mid-sized company, the firm has been growing rapidly in recent years and is expanding operations both within the United States and globally. While focused within the Midwest, the firm has operations in six different states spread widely across the United States, as well as in two different locations in China. The survey was administered to all employees in all locations, with a response rate of roughly 90%.

Overall, the company has a rather progressive reputation with regard to the treatment of their employees. The recent expansion of operations greatly increased the diversity of the workforce and helped create the interest for the current employee attitude survey. The current survey is actually a refined second wave of the survey with the first wave taking place in 2009. The most recent wave ironed out some minor issues with the first survey wave, as well as

expanded the types of questions asked. Therefore, the current study uses only the most recent wave of the survey.

The 2011 data is used to estimate a standard ordered probit. The model uses an overall measure of job satisfaction with a variety of explanatory variables as controls. The full model takes the following form:

Job Satisfaction = f(Country, Job characteristics, Context, Demographics).

The choice of how to capture job satisfaction is not without some controversy and is an ongoing issue in the job satisfaction literature. Some studies create a composite measure of satisfaction derived from responses to specific elements of the job, while others rely on a single global measure of satisfaction taken from one question. Proponents of the composite measurement, primarily in psychology and management, claim that a single measure of satisfaction cannot capture the variety of dimensions that influence a worker's satisfaction. Researchers using the single measure, primarily in economics, counter that a single measure of satisfaction can be an even more comprehensive measure of satisfaction as respondents can consider whatever elements they like when answering the satisfaction question and are not limited to the dimensions given by survey questions. In practice, studies have found that the two types of satisfaction measures may perform equally well in capturing the determinants of satisfaction (Nagy, 2002; Staples and Higgins, 1998; Wanous, Reichers, and Hudy, 1997).

The current study follows the literature that uses a single global measure of job satisfaction. The satisfaction dependent variable is derived from a question asking, "All things considered, how satisfied are you working at the company?" with seven possible responses of strongly satisfied, satisfied, somewhat satisfied, neither satisfied nor dissatisfied, somewhat dissatisfied, dissatisfied, and strongly dissatisfied. For the variable, 'strongly satisfied' is coded as the top response meaning that positive coefficients show a greater likelihood of reporting the highest level of satisfaction. With the satisfaction information derived from a single overall question, the respondent is free to consider all elements of the job, weighting various specific aspects of the job as they see fit. This measure should be influenced less by one single element of the job and does not force the researcher to impose weights on the importance of various job factors. As such the variable should be a global measure of satisfaction that captures elements that a combined measure of satisfaction might miss.

For the focus of the study examining differences across the United States and China I use a few methods to isolate possible effects. As a first check for simple differences across countries I estimate the model using only a dummy variable for the worker being in the United States. This model should capture whether there are any fundamental differences in satisfaction for Chinese workers relative to their American counterparts outside of the differences in their characteristics. If the Chinese worker is just happier on the job than the American worker the effect will show up in this variable. However, if the Chinese worker is not more or less satisfied overall than the American worker, but is influenced by different things when forming their level of satisfaction, this simple dummy will not capture the effect. Therefore, as a more detailed second check I use

location interactions, with the dummy variable for United States interacted with the other explanatory variables. The interaction terms will test for different influences on satisfaction across countries and allows a simple test of statistical significance across the two groups.

The model also exploits the detail of the survey to control for a variety of characteristics of the job and the employee. When designing the survey the company was interested in the basic level of satisfaction of workers, but was also very interested in how employees felt about more focused aspects of the job and the company such as communication, training, and morale. The depth of information allows me to investigate a range of specific determinants of employee satisfaction unavailable in other surveys. Detailed information about the variables included can be found in Table 1. After removing observations with missing values the final sample consisted of 1,719 individuals.

The second set of explanatory variables contains a number of questions asking employees about various characteristics and attitudes towards the work environment in their job. This set of questions represents a diverse range of job information and should reflect many elements of the job that drive worker satisfaction. From the information I include a variable about the level of communication in the firm, the ability to contribute to the firm as an employee, the adequacy of employee training, whether the employee feels respected by supervisors, the feeling of competitiveness of wages, the level of teamwork, and the overall morale in the company. Although the variables all utilized seven-point response scales in the initial survey, I collapse them into three categories of positive, neutral, and negative responses for easier use as explanatory variables. As all the characteristics are generally considered to be positive elements of the workplace, it is expected that positive responses to the questions should positively influence job satisfaction. The set of variables includes some influences that were found to be significant in other studies, such as training opportunities and relative wages, but also includes new variables which should expand the body of knowledge. The variables are also things that management may be able to influence making their impacts particularly interesting as they may reveal ways that management could cultivate job satisfaction in employees.

The third set of variables provides contextual information for the employee's responses. These questions ask less about the details of the company itself and more about the general employee attitude toward the company and work in general. The first variable asks about the likelihood of management action in response to the survey. This variable should give an interesting look at whether employees value the responsiveness of management or whether it is more of a secondary concern. I also include two variables asking about the development of work friendships and networking opportunities to investigate the social nature of work. All of the variables are dummies indicating a response of 'yes' to the question. In general it is expected that positive responses to the questions should positively influence satisfaction.

Table 1: Variable Names and Definitions	
Variable Name	Variable Definition
Satisfaction	All things considered, how satisfied are you working at the company? (6=Strongly satisfied, 3=Neither, 0=Strongly dissatisfied)
LOCATION	
United States	1=United States, 0=China
JOB CHARACTERISTICS	
Communication	I receive enough information so I can properly perform my job. (2=Agree, 1=Neutral, 0=Disagree)
Contribution	At work, I have the opportunity to perform my job to the best of my abilities.(2=Agree, 1=Neutral, 0=Disagree)
Training	I am adequately trained for my position. (2=Agree, 1=Neutral, 0=Disagree)
Respect	My immediate supervisor respects me and treats me fairly. (2=Agree, 1=Neutral, 0=Disagree)
Wages	I feel my wages are competitive with other companies in this area. (2=Agree, 1=Neutral, 0=Disagree)
Teamwork	The teamwork within the company is: (2=Above average, 1=Average, 0=Below average)
Morale	How would you rate the morale within the company? (2=Above average, 1=Average, 0=Below average)
CONTEXT	
Management action	I feel management will take action based on the results of this survey. (1=Yes, 0=No)
Work friendships	I have developed valuable friendships at work. (1=Yes, 0=No)
Networking opportunities	I value the opportunity to network with my co-workers. (1=Yes, 0=No)
DEMOGRAPHICS	
Employee type	1=Production, 0=Office
Years of service	1=0 to 3 years, 2=4 to 5 years, 3=Greater than 5 years
Age	1=Above 30, 0=Under 30
Supervisory status	1=Supervisor, 0=Non-supervisory

The final set of variables is comprised of demographic information used primarily as controls for the other explanatory variables. I include a dummy for being a production worker (office worker as the reference group), a set of dummies for years of service of 4 to 5 years, and

greater than 5 years (less than 3 years as the reference group), a dummy for being over 30 years old (under 30 years old the reference group), and a dummy for being a supervisor (non-supervisor as the reference group). The demographic variables included should help control for documented differences in responses across employee type and across employee characteristics, for instance, production workers tend to be less satisfied on the job than office workers. However, particularly for years of service and age, the categories had to be condensed greatly due to a lack of observations in China. Chinese workers were overwhelmingly of short tenure, partly because the operations are recently new, but also because of the high turnover in the facilities. The workers are also quite young as most are migrant workers from the interior of the country looking for opportunities on the coast. Because of the limitation a more detailed analysis could not be done with these variables. Similar issues constrained the use of more demographic variables. The survey does contain other information about the types and locations of the workers, but the lack of variation within China limited the use of the other controls. For instance, union status could not be included as all Chinese workers are non-union by definition. Unfortunately, the survey did not ask more personal information such as sex, race, etc. due to concerns about protecting the confidentiality of respondents. Therefore, analysis along the personal dimensions of the employees is not possible.

Together the sets of variables should provide valuable information about the determinants of job satisfaction, as well as controlling for potentially confounding influences on the results. However, the primary focus of the study is differences across the countries. As a simple first check for differences across countries I examine the proportion of workers responding positively to the questions for both countries. The proportion of positive responses, standard deviation, and t-test for equality are presented in Table 2. Workers in both countries report being reasonably satisfied at work with 78.1% responding positively in the U.S. and 74.0% responding positively in China, a difference that is not statistically significant. Turning to the control variables there are significant differences across countries, with U.S. workers feeling significantly more positively about their level of training and the competitiveness of their wages, but significantly less positively about the level of teamwork and worker morale. Workers in the U.S. are also significantly less likely to believe that management will take action on the survey and value networking with co-workers less than Chinese workers. The bottom panel shows clear differences in terms of worker demographics with workers in the U.S. being significantly more likely to be a production worker, of longer tenure on the job, older, and non-supervisory.

The lack of a significant difference in mean satisfaction across countries does suggest that there may not be an overall cross-country difference in level of satisfaction. The first regression model utilizing a country dummy will provide a more definitive test of this statement by controlling for worker characteristics. At the same time, the number of significant differences in job and worker characteristics across countries suggests that Chinese and American workers may well be influenced by different factors on the job. The interaction models will help investigate this thought further and will provide a more detailed view of worker attitudes in the two countries.

Table 2: United States and China, Proportion with Positive Response					
	US (N=1,670)		China (N=50)		t-stat For
Variable Names	Prop. Pos.	Std. Dev.	Prop. Pos.	Std. Dev.	Equality[a]
Satisfaction	0.781	0.413	0.740	0.443	0.653
JOB CHARACTERISTICS					
Communication	0.743	0.437	0.800	0.404	0.979
Contribution	0.821	0.384	0.840	0.370	0.358
Training	0.862	0.345	0.420	0.499	6.220***
Respect	0.793	0.405	0.760	0.431	0.541
Wages	0.613	0.487	0.340	0.479	3.976***
Teamwork	0.509	0.500	0.640	0.485	1.881*
Morale	0.405	0.491	0.680	0.471	4.064***
CONTEXT					
Management action	0.523	0.500	0.820	0.388	5.276***
Work friendships	0.849	0.358	0.900	0.303	1.164
Networking opportunities	0.804	0.397	0.940	0.240	3.848***
DEMOGRAPHICS					
Employee type: Production	0.674	0.469	0.440	0.501	3.261***
Years of service: % 0-3 yrs	0.200	-	0.400	-	Chi2=45.89***
Years of service: % 4-5 yrs	0.173	-	0.440	-	
Years of service: % >5 yrs	0.627	-	0.160	-	
Age: Above 30	0.870	0.336	0.600	0.495	3.832***
Supervisory status	0.122	0.327	0.260	0.443	2.192**
***Statistically significant at the 1% level **at the 5% level *at the 10% level (two tailed tests).					
[a] For Years of Service the entire distribution is examined, with a Chi Square test of the equality of the distributions used instead of a t-test.					

RESULTS

The results for the baseline model using only a country dummy variable are presented in the first column of Table 3. The table contains the marginal effect of each variable on reporting the highest level of satisfaction (strongly satisfied). The marginal effects are the most intuitive way of understanding the results of an ordered probit model as the coefficients themselves can be confusing and even misleading. Not only are the magnitudes of the coefficients hard to interpret, but the signs can even be opposite of the marginal effects. To simplify the analysis of the results I present only the marginal effect on the highest category of satisfaction.

The first thing to note in the table is that there is not any significant overall difference in satisfaction across countries. The dummy variable for being in the United States is insignificant after controlling for job and worker differences. Chinese workers do not appear to be any more, or less, satisfied than their American counterparts. This result matches the simple evidence provided with the insignificant difference in means in Table 2. As mentioned previously,

however, the result does not show whether the level of satisfaction is determined by different factors in each country. The interaction model is needed to further investigate this question.

Table 3: Ordered Probit Marginal Effects on Highest Category of Satisfaction, Various Specifications

	Indicator Model	Interaction Model		
	Both Countries	United States	China	Z Stat for Equality Across Countries
Variable Names	P(Strongly sat.)	P(Strongly sat.)	P(Strongly sat.)	
United States	0.033	-	-	-
JOB CHARS.				
Communication:				
Neutral	0.014	0.011	0.415**	-2.31**
Agree	0.028**	0.023*	0.161***	-1.84*
Contribution:				
Neutral	-0.020	-0.015	-0.238	0.99
Agree	0.025*	0.029**	-0.092	0.78
Training:				
Neutral	0.014	0.021	-0.216*	1.97**
Agree	0.025	0.034**	-0.242*	2.28**
Respect:				
Neutral	0.024	0.028*	0.193	-0.92
Agree	0.067***	0.070***	0.035	0.25
Wages:				
Neutral	0.045***	0.046***	0.064	-0.31
Agree	0.115***	0.111***	0.268***	-1.79*
Teamwork:				
Average	0.024**	0.024**	0.058	-0.14
Above Average	0.045***	0.047***	-0.161	1.90*
Morale:				
Average	0.046***	0.045***	0.071	-1.00
Above Average	0.151***	0.149***	0.287***	-2.49**
CONTEXT				
Man. Action	0.054***	0.058***	-0.263	2.37**
Work Friendships	0.042***	0.042***	0.160***	-1.99**
Networking Opps.	0.002	0.004	-0.300	1.82*
DEMOGRAPHICS				
Production Emp.	-0.002	-0.002	-0.066	0.89
Years of Service:				
4 to 5 Years	0.003	0.003	-0.137**	2.24**
> than 5 years	0.006	0.002	-0.003	0.05
Above 30	0.019	0.032**	-0.229***	3.54***
Supervisor	0.022	0.021	0.089	-0.77
	N=1,720	N=1,670	N=50	

***Statistically significant at the 1% level **at the 5% level *at the 10% level (two tailed tests).

Before turning to the interaction model it is useful to look at results for the control variables as a number of significant predictors of worker satisfaction were found. In the job characteristics group of controls, those workers who agree with the statement that they receive enough information to perform their job are 2.8 percentage points more likely to report the highest level of satisfaction, while those who agree that their supervisor respects them are 6.7 points more likely to be strongly satisfied. Those who feel that teamwork is above average or average also are more likely to be strongly satisfied, with positive boosts of 4.5 and 2.4 percentage points respectively. Feelings about wages and morale have even greater influences on worker satisfaction. Workers who agree or are neutral about the statement that their wages are competitive experience increases in the likelihood of being strongly satisfied of 11.5 and 4.5 percentage points respectively. The fact that both positive and neutral feelings about wage competitiveness increase satisfaction shows how strong a negative influence the perception of uncompetitive wages can be in the work place. These results match the general findings in previous work about the effect of relative wages. The perception component of this result is important to note as well as it may not be that wages truly are not competitive in the firm, but simply that they are being perceived as not competitive. If management knows that wages truly do match wages in surrounding firms, it may help to communicate the evidence. If the perception of uncompetitive wages changes it might increase satisfaction levels without actually having to raise the monetary wages. A similar pattern is observed with morale, as those who feel that morale is above average, 15.1 points, and those who feel it is just average, 4.6 points, both see significant increases in the likelihood of being strongly satisfied. There is a marginally significant, and small, positive effect for workers who agree that they have the opportunity to perform their job to the best of their abilities and no significant effect of being adequately trained. The significant results generally match expectations.

For the context variables, those who feel that management will take action based on the survey and who developed work friendships see significant increases in satisfaction of 5.4 and 4.2 percentage points respectively. The results are interesting as they suggest a few paths that management could take to increase satisfaction. One method would be to appear more responsive to worker concerns. Firms have tried various methods to implement greater employee participation at work in both union and non-union settings. The results suggest that workers appreciate these efforts and that firms who have not implemented mechanisms for employee input might benefit from doing so, even outside of the actual input received. Another method is to cultivate a company atmosphere that allows workers to socialize with each other. This finding does not necessarily mean that everyone should have a company picnic, but it suggests that even small changes like a dedicated break room might have beneficial effects on overall satisfaction if they help workers make friends with coworkers.

Interestingly, none of the demographic variables are significant predictors of being highly satisfied. This result suggests that any differences seen across demographic groups should really be attributed to differences in job characteristics across groups and not really to the people in the

groups themselves. For example, if it is found that younger workers are more satisfied, it may not really be due to something about youth, but rather that young workers are just more likely to feel that morale is above average in the company. Once you take morale out of the comparison, young and old workers are essentially the same. This observation is important as attitudes of workers can sometimes be changed, while demographic characteristics such as age are usually outside of the control of management.

The extension of the baseline indicator model incorporating country interactions will provide a more detailed picture of possible differences in satisfaction across countries. While the baseline model suggests there is no overall difference in satisfaction across workers in different countries, it cannot answer the question of whether those similar levels of satisfaction are driven by different factors. The results from the model interacting each explanatory variable with the U.S. dummy variable are presented in the right panel of Table 3. Once again the marginal effects on the probability of being Strongly Satisfied are reported. Although the model is estimated jointly, the marginal effect of each variable is calculated separately for each country. The results for the United States are presented in the first column while the results for China are presented in the second column. The third column contains the Z-stat for equality of coefficients across countries taken from the interaction term in the ordered probit regression model.

Examining the results for workers in the U.S. in the first column, we can see that the results generally match those in the indicator model. The similarity of the results is not overly surprising as American workers do form the vast majority of the observations in the indicator model. The only notable differences with the indicator model are that those who agree that they are adequately trained are 3.4 percentage points more likely to be strongly satisfied, and workers above 30 are 3.2 percentage points more likely to be strongly satisfied. These differences are notable though as they are two results found in prior work that did not show up in the indicator model combining both countries.

Turning to the results for workers in China some interesting differences become evident across countries. For the job characteristics questions in the top panel, agreeing that communication is good has a significant positive effect in China as in the United States, but the magnitude of the effect is much greater in China, 16.1 points compared to only 2.3 points, and is in fact statistically different than the response in the United States. Additionally, workers in China who are just neutral about the communication question see a large positive effect of 41.5 percentage points on the probability of being strongly satisfied, a significantly larger effect than in the United States. A similar result can be seen with morale as rating morale above average increases the likelihood of being strongly satisfied by 28.7 points in China but only 14.9 points in the United States, a difference that is statistically significant once again. There is also some evidence that Chinese workers are more positively influenced by perceptions that wages are competitive, with workers who agree in China experiencing a 26.8 percentage point increase in the likelihood of being strongly satisfied compared to only 11.1 points in the United States. This difference is marginally significant across countries. The result for communication is important as it suggests that managers should have different strategies for their Chinese and American

plants if they are looking to increase worker satisfaction. While good communication is valued everywhere to some extent, improving methods of communication will only have modest returns in the U.S., but may greatly increase satisfaction in China. This may also be a relatively cost effective way to improve satisfaction in China as distributing information is likely less expensive than paying overall higher wages or other methods of raising satisfaction. The results also suggest that efforts to boost morale in China could bring a great return if successful.

Chinese workers also have stronger negative reactions than American workers to a few variables. The most interesting result is for the statement that the worker is adequately trained. In the indicator model there is no significant effect for the training variable. However, in China there are marginally significant, strongly negative reactions for those who agree to or are neutral to the question that they are adequately trained. Those workers in China who agree that they are adequately trained experience a 24.2 point reduction in the likelihood of being strongly satisfied, while those that are neutral to the question experience a 21.6 point reduction in the likelihood. These differences are both highly significant across countries. On the surface this reaction to training may seem counterintuitive and contrary to prior research. However, the explanation for the negative reaction could be due to the types of jobs within the company allocated between the U.S. and China. If jobs sent to China by U.S. companies are overall lower skill, the average job in China will require much less training than in the United States. The Chinese worker responding that they are adequately trained for their job might not really be reacting negatively to the training, but rather to the fact that the job they do requires little skill and is in general boring or alienating. If this is the case, managers likely cannot do much with the training process to alleviate the negative effects on satisfaction, but should be aware of them so that they can possibly be offset in another manner. Additionally, responding that teamwork is above average has a marginally significant different effect across countries, despite the fact that the teamwork rating is not significant within China. The signs and significance of the individual country coefficients, positive and significant in the United States and insignificant in China, suggest a negative effect for Chinese workers relative to American workers, but the low level of statistical significance across countries suggests the result should not be interpreted too strongly.

There are also significant differences across countries for the context variables. Chinese workers reacted significantly more positively than American workers to the formation of work friendships, with an increase in the likelihood of being strongly satisfied of 16.0 points in China and only 4.2 points in the United States. This reaction may be attributable to the types of workers who take the jobs in China. Overwhelmingly the workers in the Chinese plants are migrant workers from the interior of the country who have moved to the coast for factory jobs. These workers are generally housed in dormitories and spend virtually their entire time in and around the plant. For these workers, friendships formed on the job may be vital to satisfaction as they make up a high percentage of their overall social interaction and friendships. In the United States the effect of work friendships may be muted because there is more separation between work and leisure. An American worker does appreciate work friendships, but when they leave the plant

they may have many other friends that have no connection with the job, making the work friendships less vital to overall satisfaction. This interpretation might be somewhat supported by the greater positive reaction to morale for Chinese workers highlighted in the job characteristics. Since the Chinese worker has a harder time separating from the job and fellow workers, the more positive reaction to morale may reflect a greater importance of social interaction on the job. If this is true, companies looking to increase satisfaction of Chinese workers in these types of settings should focus efforts on worker interaction and socialization. The social aspects of the dormitories and non-work life may also have spillover effects on satisfaction on the job.

An interesting negative difference for Chinese workers arises in regard to perceptions of whether management would take action based on the survey. Similar to the indicator model estimates, workers in the United States perceive management action positively, with an increase in satisfaction for those who thought management would act based on the results. However, in China the reaction to perceptions of management action is insignificant, with the difference across countries highly significant. The divergent results suggest that American workers expect management action to be proactive and positive, while Chinese workers expect any management action to be disciplinary or negative in nature. The reaction may be cultural and is important for managers in American companies to note. Even if the intentions of plant managers in China are benevolent, simply offering an employee survey to workers may have a strong negative effect on satisfaction if workers expect to be disciplined as a result. American firms looking to use employee participation or input strategies in China will need to take strong actions to try and reassure employees that there will not be a harsh management reaction to their responses. If this perception of management retribution cannot be changed, surveys like this may actually cause more harm than benefit. There is some evidence for a differential reaction to networking opportunities as well with a marginally significant negative difference found for Chinese workers relative to their counterparts in the United States. However, neither of the coefficients within countries is significantly different from zero casting doubt on how strong the effect can actually be.

In terms of the demographic variables there are some interesting differences as well. Satisfaction for Chinese workers declines as they age with workers above 30 experiencing a 22.9 point reduction in the likelihood of high satisfaction, while American workers above 30 actually experience a 3.2 point increase in likelihood. A similar result may be evident for years of service with Chinese workers who have been with the company 4-5 years experiencing a 13.7 point reduction in the likelihood of being strongly satisfied. These results may be due to a fundamental difference in reaction to aging for Chinese workers, but it also may be due to the unusual characteristics of the migrant work force. The migrant workforce tends to be young with workers returning to their homes as they age. Some of the motivation to return home may be due to elderly relatives but a great deal is also due to the Chinese hukou system of household registration. Workers who have moved to the cities who have a rural hukou registration are not able to access basic social services in the cities. For young workers looking to start a family this is especially problematic as their children will not have access to public education unless they

return home. For many young migrants the move from home can be viewed as a temporary strategy in order to accumulate resources for a life at home. In this context, the negative responses for workers over 30 who are still working in the jobs may reflect a failure to make the temporary nature of the move a reality, and may reflect dissatisfaction with their life in general rather than dissatisfaction with the job per se. Similarly the migrant workforce is very fluid, with high turnover on jobs. Part of the turnover is workers moving back home, but part of the turnover is workers transitioning between jobs in the area looking for better opportunities. In fact, one study finds that this job mobility can help migrants eliminate their observed wage disadvantage in urban labor markets (Ariga et al., 2012). In this context, workers who have been on the job 4-5 years may be dissatisfied because they have not been able to find better opportunities and may feel stuck in their current position. Once again, this may say less about the job per se, and more about a generally frustrated worker.

Despite the number of differences across countries there are a number of similarities as well. Within the United States workers have a positive reaction to feeling respected and a positive reaction to being able to contribute, both of which are insignificant within China. However, the differences for both variables are insignificant across countries. There are also no significant differences for production or supervisory workers.

The results from the interaction model show some significant and interesting differences across countries that can have practical importance for companies with operations in both the United States and China. To check the robustness of the findings I perform a few specification checks to see if the results are influenced by the form of the model. A first concern relates to the possible presence of multicollinearity in the explanatory variables. Given the interrelated nature of a job, one could reasonably expect that attitudes towards portions of the job may be strongly related. If the explanatory variables are too closely related severe multicollinearity could be driving the results. Unfortunately multicollinearity is a difficult issue to pin down directly and the determination of whether the issue is a problem is somewhat subjective. As a first check I conduct simple tests for multicollinearity which do not suggest any problems at all. The variance inflation factors are all below standard, acceptable levels. However, as a second check I estimate the model without the morale variable to see if the pattern of results changes dramatically. The motivation for excluding morale specifically is that the question is more global in nature than the other, more specific questions about the job, and therefore, may be closely related to each of the individual questions. It is possible that if one rates morale within the company highly that they also are very likely to rate the other components highly.

Results from the specification test excluding morale are presented in the first three columns of Table 4. Examining the results the first thing to notice is that the findings for the United States are virtually identical to the baseline estimates in Table 3. The magnitudes of the marginal effects and significance levels are extremely similar suggesting that multicollinearity

Table 4: Ordered Probit Marginal Effects on Highest Category of Satisfaction, by Country						
	No Morale			U.S. Non-Union		
Variable Names	U.S. P(Str. sat.)	China P(Str. sat.)	Z Stat for Equality	U.S. P(Str. sat.)	China P(Str. sat.)	Z Stat for Equality
JOB CHARS.						
Communication:						
Neutral	0.021	0.219	-1.21	0.077**	0.400**	-2.09**
Agree	0.031**	0.103	-0.86	0.062**	0.152***	-1.66*
Contribution:						
Neutral	-0.009	-0.070	0.25	0.022	-0.228	1.15
Agree	0.044***	0.001	0.29	0.089***	-0.092	1.23
Training:						
Neutral	0.018	0.004	0.15	0.055	-0.213*	2.18**
Agree	0.034**	-0.026	0.64	0.050*	-0.237*	2.41**
Respect:						
Neutral	0.024	0.023	-0.07	0.020	0.190	-1.05
Agree	0.073***	0.110	-0.49	0.074***	0.035	0.18
Wages:						
Neutral	0.045***	0.021	0.36	0.044**	0.059	-0.55
Agree	0.119***	0.254***	-1.26	0.106***	0.263***	-2.36**
Teamwork:						
Average	0.037***	0.028	0.13	-0.006	0.050	-0.32
Above Average	0.097***	0.009	0.96	0.045*	-0.162	1.88*
Morale:						
Average	-	-	-	0.041***	0.060	-1.02
Above Average	-	-	-	0.162***	0.272***	-2.50**
CONTEXT						
Man. Action	0.092***	-0.352**	3.18***	0.057***	-0.268*	2.47**
Work Friendships	0.043***	0.138***	-1.34	0.066***	0.148***	-1.98**
Networking Opps.	0.011	-0.105	0.78	0.013	-0.281	1.86*
DEMOGRAPHICS						
Production Emp.	-0.008	-0.027	0.25	-0.058***	-0.062	0.30
Years of Service:						
4 to 5 Years	-0.008	-0.062	0.99	-0.041*	-0.133**	1.78*
> than 5 years	-0.013	0.066	-0.78	-0.018	-0.002	-0.13
Above 30	0.031**	-0.245***	3.74***	0.061***	-0.223***	3.96***
Supervisor	0.017	0.083	-0.68	0.031	0.085	-0.75
	N=1,675	N=50		N=694	N=50	
***Statistically significant at the 1% level **at the 5% level *at the 10% level (two tailed tests).						

with the morale variable is not a great issue. However, there are some differences for the smaller sample China results and the cross country significance levels. The communication responses and 4 to 5 years of service response lose their within country significance, as do the previously marginally significant training responses. All of these variables also lose their cross country

significance. The changes suggest that multicollinearity may be influencing the Chinese estimates to some extent. At the same time there are similar results for wages, management action, work friendships, and being above 30, suggesting that even if multicollinearity is an issue it is not driving all of the findings.

Despite the fluctuations there are good reasons to believe that multicollinearity is not a major problem in the model. First, while one symptom of multicollinearity is results changing due to small changes in the model, typically the individual coefficient estimates gain significance when the offending variable is removed rather than lose significance. This pattern is due to the fact that multicollinearity increases coefficient standard errors leading to low significance levels. This is the opposite of what happened in this situation. Second, and perhaps most importantly, if the morale variable truly does contain separate explanatory information from the other individual questions, removing the predictive variable from the model will causes changes in the estimates as well. The estimates of the more specific question coefficients will be influenced by the omitted morale effect. I would argue that this is the case here. While the morale question is more global in nature than the other more specific questions, rating worker morale is still a much different thing than rating the effectiveness of communication or the adequacy of training. Feelings of morale are likely to influence response to the other questions, but likely not to the extent that they are completely driving the results. For these reasons, combined with the fact that the simple tests showed no influence of multicollinearity, I believe that the baseline findings are not driven by multicollinearity.

A second specification check relates to the presence of unions in some of the U.S. plants. In the United States, the production workers are union in some states and non-union in others, while they are all non-union in China. This is problematic for the results as there is a systematic difference across workers in the two countries that I cannot control for adequately given the lack of variation in union status in China. If union membership does influence satisfaction, as suggested by prior work, this causes a difference in the control and treatment group other than country of the worker, potentially biasing the results. Although I cannot control for the issue using an included union variable, I can check for a possible effect of unions by excluding U.S. union workers from the sample and estimating the model. The exclusion of union workers will make the Chinese worker more comparable to their American counterparts.

The results for the non-union specification are presented in the last three columns of Table 4. The most significant thing to note in regard to the results is that they are very similar to the results for the baseline model. There are no changes in the China estimates with the exception that the negative coefficient on perceived management action now becomes marginally significant. There are also virtually no changes in either direction or significance level for the cross country differences. As could be expected given that it was the United States portion of the sample that was influenced, there are some small changes in the United States results. For non-union workers in the United States communication is more highly valued, perhaps because the union is not occupying the role of gathering and distributing information to

workers. In addition, non-union production workers now experience a 5.8 percentage point reduction in the likelihood of being strongly satisfied when there was no significant effect previously. Having 4 to 5 years of service now also reduces the likelihood of being satisfied by a marginally significant 4.1 percentage points. This effect likely reflects the fact that for non-union workers tenure on the job does not bring as many perks as in a union setting where seniority likely plays a much greater role. While there are very minor differences when union members are excluded from the sample, the results suggest that the presence of union members in the United States is not driving the primary findings. In fact, the results in Table 4 may actually be more accurate estimates. Chinese workers react more positively than American workers to effective communication, competitive wages, good morale, and work friendships. On the other hand, Chinese workers react more negatively than American workers to being adequately trained, perceptions of management action, longer tenure on the job, and being above 30.

Of the results established in the baseline and non-union models presented in Tables 3 and 4, one that deserves a bit more attention is the seemingly counterintuitive result for perceptions of adequate training in China. One explanation put forth is that the jobs in China may simply be lower skill jobs which require less training. Workers may respond that they are adequately trained, but rather than this being an indication of satisfaction from having obtained a higher level of skill as in the United States, Chinese workers may feel that their adequate training simply reflects little opportunity to actually develop skills or to advance in the organization. To try and disentangle these possible effects I estimate the model including two questions about the development of skills and promotion opportunities. The first question asks, "In the last year, I have had opportunities to learn and develop new skills at work.", while the second asks, "Opportunities exist for promotion within the company." Respondents are asked to agree or disagree with the questions on a 7 point scale. If skill development and promotion opportunities are wrapped up in the training question estimates the inclusion of these two variables should help separate the effects.

Results for this specification are presented in Table 5, with the first three columns repeating the baseline estimates for ease of comparison, and the last three columns presenting the model including the skill development and promotion variables. The first conclusion that can be drawn is that there is a great deal of stability between the models. Overall the within country marginal effects and the cross country differences are very similar with only marginal changes in significance. The only real exception is for the communication variable where the positive effect for China is somewhat reduced. The second notable point is that workers do react as expected to skill development and promotion opportunities, but only in the United States. Workers in the United States who agree that they have developed skills within the last year are 2.9 percentage points more likely to be strongly satisfied, while those who agree that promotion opportunities exist are 7.3 percentage points more likely to be strongly satisfied. The variables are insignificant within China, although this should probably not be interpreted too strongly as there are also insignificant differences across the two countries.

Table 5: Ordered Probit Marginal Effects on Highest Category of Satisfaction, by Country						
	Baseline Model			Training and Advancement Model		
Variable Names	U.S. P(Str. sat.)	China P(Str. sat.)	Z Stat for Equality	U.S. P(Str. sat.)	China P(Str. sat.)	Z Stat for Equality
JOB CHARS.						
Communication:						
Neutral	0.011	0.415**	-2.31**	0.013	0.377*	-1.96*
Agree	0.023*	0.161***	-1.84*	0.018	0.132*	-1.29
Contribution:						
Neutral	-0.015	-0.238	0.99	-0.015	-0.228	1.02
Agree	0.029**	-0.092	0.78	0.028**	-0.057	0.54
Training:						
Neutral	0.021	-0.216*	1.97**	0.012	-0.239*	1.95*
Agree	0.034**	-0.242*	2.28**	0.030**	-0.376***	2.85***
Respect:						
Neutral	0.028*	0.193	-0.92	0.024	0.076	-0.27
Agree	0.070***	0.035	0.25	0.063***	-0.006	0.50
Wages:						
Neutral	0.046***	0.064	-0.31	0.040***	0.042	0.00
Agree	0.111***	0.268***	-1.79*	0.098***	0.165*	-0.71
Teamwork:						
Average	0.024**	0.058	-0.14	0.020	0.284	-1.29
Above Average	0.047***	-0.161	1.90*	0.038***	-0.125	1.57
Morale:						
Average	0.045***	0.071	-1.00	0.043***	0.029	-0.21
Above Average	0.149***	0.287***	-2.49**	0.134***	0.278***	-2.07**
Develop skills:						
Neutral	-	-	-	0.003	-0.061	1.01
Agree	-	-	-	0.029**	0.156	-1.17
Promotion Opps:						
Neutral	-	-	-	0.041***	-0.045	1.02
Agree	-	-	-	0.073***	0.064	0.06
CONTEXT						
Man. Action	0.058***	-0.263	2.37**	0.048***	-0.197	1.81*
Work Friendships	0.042***	0.160***	-1.99**	0.039***	0.168***	-2.01**
Networking Opps.	0.004	-0.300	1.82*	0.003	-0.258	1.57
DEMOGRAPHICS						
Production Emp.	-0.002	-0.066	0.89	0.001	-0.081	1.18
Years of Service:						
4 to 5 Years	0.003	-0.137**	2.24**	0.010	-0.121*	2.00**
> than 5 years	0.002	-0.003	0.05	0.010	-0.026	0.34
Above 30	0.032**	-0.229***	3.54***	0.030**	-0.154*	2.53**
Supervisor	0.021	0.089	-0.77	0.013	0.065	-0.60
	N=1,670	N=50		N=1,660	N=50	
***Statistically significant at the 1% level **at the 5% level *at the 10% level (two tailed tests).						

Of most importance however is the fact that the inclusion of the two variables does nothing to alleviate the negative effect of adequate training in China, and actually increases the size and significance of the effect to some extent. This negative training result may still be due to the potentially alienating nature of low skill work, but future research could investigate the topic in more detail.

CONCLUSION

With the rapid expansion of operations in China companies are finding it is increasingly important, and difficult, to understand their workforce. This study examines the job satisfaction of workers in the United States and China, testing whether there are different influences on satisfaction across countries. To conduct the study I exploit a detailed dataset from a regional manufacturing firm with operations in both countries. I use the data to estimate an ordered probit with a global measure of job satisfaction as the dependent variable. I interact a country dummy with each explanatory variable to allow for differential effects across countries and to facilitate a statistical test of differences across countries.

While there does not appear to be an overall difference across countries in the level of job satisfaction, the job satisfaction in each country does appear to be driven by different factors. Chinese workers respond more positively to good communication than American workers. Chinese workers also seem to value the communal atmosphere at work more than American workers as they respond more positively to both morale and work friendships. This difference may be due to the largely migrant workforce in China which experiences less separation between work and social life than do workers in the United States. In addition, although workers everywhere seem to value money, the positive effect of perceptions of competitive wages is stronger for Chinese workers than for their American counterparts. On the other hand, workers in China have strong negative reactions to feeling adequately training as opposed to positive feelings in the United States, perhaps due to the fact that the jobs sent to China require less skill overall and lead to boring, monotonous work. Chinese workers also seem to fear negative management actions as there is a strong negative effect of the likelihood of management action based on the survey compared to a positive reaction for American workers. This finding is important to note as American companies may need to reassure their Chinese workers of their intentions, or efforts to interact with workers may cause more damage than benefit. Older Chinese workers and those with longer tenures also seem to be less satisfied compared to their American equivalents, perhaps due to frustration that perceived opportunities on the job did not turn out the way they planned.

The findings are important as they suggest different strategies to boost worker satisfaction across countries for American firms with Chinese operations. Some strategies may have similar effects in both countries, while others may have stronger effects in one country, or may work in only one. If American companies are to get the most out of their operations in

China, knowing and exploiting these cross country differentials is essential. As the Chinese labor market tightens, satisfying the Chinese worker may be necessary to attract qualified labor.

REFERENCES

Ariga, Kenn, Fumio Ohtake, Masaru Sasaki, and Zheren Wu (2012). Wage Growth Through Job Hopping in China. *IZA Discussion Paper*, No. 7104.

Artz, Benjamin (2008). The Role of Firm Size and Performance Pay in Determining Employee Job Satisfaction. *Labour,* 22(2), 315-43.

Bockerman, Petri, and Pekka Ilmakunnas (2009). Job Disamenities, Job Satisfaction, Quit Intentions, and Actual Separations: Putting the Pieces Together. *Industrial Relations,* 48(1), 73-96.

Brown, Gordon D. A., Jonathan Gardner, Andrew J. Oswald, and Jing Qian (2008). Does Wage Rank Affect Employees' Well-being? *Industrial Relations,* 47(3), 355-89.

Clark, Andrew E., Nicolai Kristensen, and Niels Westergard-Nielsen (2009). Job Satisfaction and Co-Worker Wages: Status or Signal? *The Economic Journal,* 119(March), 430-47.

Donald, Ian and Ol-Ling Siu (2001). Moderating the Stress Impact of Environmental Conditions: The Effect of Organizational Commitment in Hong Kong and China. *Journal of Environmental Psychology,* 21(4), 353-68.

Drago, Robert, and Mark Wooden (1992). The Determinants of Labor Absence: Economic Factors and Workgroup Norms across Countries. *Industrial and Labor Relations Review,* 45(4), 764-78.

Gao, Wenshu and Russell Smyth (2009). Job Satisfaction and Relative Income in Economic Transition: Status or Signal? The Case of Urban China. *Monash University, Development Research Unit Discussion Paper,* DEVDP 09-12.

Gazioglu, Saziye and Aysit Tansel (2006). Job Satisfaction in Britain: Individual and Job Related Factors. *Applied Economics,* 38, 1163-1171.

Green, Francis (2010). Well-Being, Job Satisfaction, and Labour Mobility. *Labour Economics*, 17(6), 897-903.

Han, Yong and Nada K. Kakabadse (2009). Job Satisfaction: What is its True Meaning in Greater China? *Asia-Pacific Journal of Business,* 1(2), 155-64.

Heywood, John S., W.S. Siebert, and Xiangdong Wei (2002). Worker Sorting and Job Satisfaction: The Case of Union and Government Jobs. *Industrial and Labor Relations Review,* 55(4), 595-609.

Heywood, John S. and Xiangdong Wei (2006). Performance Pay and Job Satisfaction. *Journal of Industrial Relations,* 48(4), 523-40.

Kosteas, Vasilios D (2011). Job Satisfaction and Promotions. *Industrial Relations,* 50(1), 174-94.

Kristensen, Nicolai and Edvard Johansson (2008). New Evidence on Cross-Country Differences in Job Satisfaction Using Anchoring Vignettes. *Labour Economics,* 15(1), 96-117.

Leung, Kwok, Zhongming Wang, and Peter B. Smith (2001). Job Attitudes and Organizational Justice in Joint Venture Hotels in China: The Role of Expatriate Managers. *International Journal of Human Resource Management,* 12(6), 926-45.

Linz, Susan J. and Anastasia Semykina (2012). What Makes Workers Happy? Anticipated Rewards and Job Satisfaction. *Industrial Relations,* 51(4), 811-844.

Ma, Shaozhuang and Virginia Trigo (2008). Winning the War for Managerial Talent in China. *The Chinese Economy,* 41(3), 34-57.

Mangione, Thomas W. and Robert P. Quinn (1975). Job Satisfaction, Counterproductive Behavior and Drug Use at Work. *Journal of Applied Psychology,* 60(1), 114-16.

Munyon, Timothy, Wayne Hochwarter, Pamela Perrewe, and Gerald Ferris (2010). Optimism and the Non-linear Citizenship Behavior-Job Satisfaction Relationship in Three Studies. *Journal of Management,* 36(6), 1505-28.

Nagy, Mark S (2002). Using a Single-Item Approach to Measure Facet Job Satisfaction. *Journal of Occupational and Organizational Psychology,* 75, 77-86.

Nielsen, Ingrid and Russell Smyth (2008). Job Satisfaction and Response to Incentives Among China's Urban Workforce. *Journal of Socio-Economics,* 37, 1931-36.

Pouliakas, Konstantinos and IoannisTheodossiou (2010). Differences in the Job Satisfaction of High-paid and Low-paid Workers Across Europe. *International Labour Review,* 149(1), 1-29.

Rutherford, Brian N., Yujie Wei, JungKun Park, and Won-Moo Hur (2012). Increasing Job Performance and Reducing Turnover: An Examination of Female Chinese Salespeople. *Journal of Marketing Theory and Practice,* 20(4), 423-36.

Siu, Ol-Ling (2002). Occupational Stressor and Well-Being Among Chinese Employees: The Role of Organizational Commitment. *Applied Psychology: An International Review,* 51(4), 527-44.

Sousa-Poza, Alfonso and Andres A. Sousa-Poza (2000). Well-being at Work: A Cross-National Analysis of the Levels and Determinants of Job Satisfaction. *Journal of Socio-Economics,* 29, 517-38.

Staples, D. Sandy and Christopher A. Higgins (1998). A Study of the Impact of Factor Importance Weightings on Job Satisfaction Measures. *Journal of Business and Psychology,* 13(2),211-32.

Wang, Huashu, Lei Pan, and Nico Heerink (2013). Work Conditions and Job Satisfaction of China's New Generation of Migrant Workers: Evidence from an Inland City. *IZA Discussion Paper*, No. 7405.

Wanous, John P., Arnon E. Reichers, and Michael J. Hudy (1997). Overall Job Satisfaction: How Good are Single-Item Measures? *Journal of Applied Psychology,* 82(2), 247-52.

Wong, Chi-Sum, Yui-tim Wong, Chun Hui, and Kenneth S. Law (2001). The Significant Role of Chinese Employees' Organizational Commitment: Implications for Managing Employees in Chinese Societies. *Journal of World Business,* 36(3), 326-40.

MODELING AFRICA'S ECONOMIC GROWTH

Oluremi Ogun, University of Ibadan,

ABSTRACT

This study investigated Africa's economic growth over the period 1996 to 2010, deriving its motivation from the theoretical and empirical literature on the subject. Factors peculiar to the continent such as conflicts, policy distortions, weak institutions, export reliance and low productivity growth were accommodated. Departure from the conventional model specification occurred in the areas of exclusion of some traditional factors and the emergence of new entries. The methodology consisted of both a static and a dynamic panel data analyses of fifteen countries distributed across the different regions of the continent. Some remarkable results were obtained.

INTRODUCTION

Over the last decades, economic growth and its determinants have been of great importance in both theoretical and applied studies. This is due to much importance of economic growth itself. The first steps towards developing the theories of economic growth were taken in the 1930's and early 1940's. All these theories have been directed to the two central questions: why growth rates across countries are different and what factors cause this difference? This difference manifests itself in different standards of living and quality of life in all over the world. In some economies, the level of investment and the productivity is low; the workers face little change in their standards of living and the growth rate and level of development are low; whereas in some other countries, these indices are high enough.

Africa's poor economic performance has been widely studied. Within the empirical growth literature, considerable attention has been paid to slow growth performance in Africa. In average term, the growth rate in Africa hardly surpassed 2% while East and the Pacific countries had over 5% and Latin America experienced growth rate above 2%n (Easterly and Levine, 1997). Large body of studies points to a diverse set of potential causes of Africa's growth tragedy, ranging from bad policies, to poor education, political instability and inadequate infrastructure, but prominent among the cause is low factor productivity growth (see Ndulu and O'Connell, 1999; Ndulu and O'Connell, 2009, Berthelemy and Söderling, 2001; Hoeffler, 2002 and Fosu, 2002). This literature has improved our understanding of African growth tragedy. However, it fails to guide us directly to the factors behind the low productivity growth observed in Africa.

This study did not directly approach the issue of determinants of productivity growth in Africa; rather, it provided quantitative estimates of the extent to which observed productivity

growth accounted for the growth performance of the continent. In so doing, it specified a growth model derived from economic theory but somewhat departed from the approach common to the literature. The rest of the paper is organized as follows. In section II, an overview of growth theory is presented succeeding in a following subsection, a review of the empirical growth literature as applied to Africa. Section III specified the empirical model employed in the study while, section IV dealt with the results and their interpretations. The final section provided some concluding remarks.

THE GROWTH LITERATURE

Overview

Classical economists, such as Smith (1991), Malthus (1798), Richardo and Eck (1817) and much later Ramsey (1928),Young (1928), Schumpeter (1934) and Knight (1944) provided many of the basic ingredients that appear in modern theories of economic growth. The main studies begin on these basic ingredients and focuses on the contributions in the neoclassical tradition since the late 1950s. From a chronological viewpoint, the starting point for modern theory growth theory is the classic article of Ramsey (1928). Ramsey's treatment of household optimization over time goes far beyond its application to growth theory. Between Ramsey and late 1950s, Harrod (1939) and Domar (1946) attempted to integrate Keynesian analysis with elements of economic growth. They used production function with little substitutability among the inputs to argue that the capitalist system is inherently unstable.

The next and more important contributions of modern growth theory have been the works of Solow (1956) and Swan (1956). The fundamental features of the Solow-Swan neoclassical production function are the assumptions of constant returns to scale, diminishing returns to each input and some positive and smooth elasticity of substitution between the inputs. The Solow-Swan production function is applied along with a constant saving rate rule in order to generate a simple general equilibrium model of the economy. A key prediction of these neoclassical growth models which has been frequently applied as an empirical hypothesis is conditional convergence, in the sense that the lower the starting level of per capita GDP, compare to the long-run or steady state position, the faster the growth rate. This is due to the assumption of diminishing returns to capital.

In the late of 1950s and 1960s, the neoclassical growth theorists came to recognize the deficiencies in the past models. In order to overcome this, these theories tend to assume that technological progress occurred in an exogenous manner. This assumption would permit a positive constant per capita long term growth rate, while retaining the prediction of conditional convergence.

Cass (1965) and Koopmans (1965) applied Ramsey's analysis of consumer optimization to the neoclassical growth model in order to make adequate preparation for an endogenous determination of the saving rate. This extension tends to preserve the hypothesis of conditional convergence, while allowing for strong transitional dynamics. Due to the lack of relevance and empirical supporting evidence, growth theory effectively came to the end as an active research

field by the early 1970s. The years after the mid-1980, have witnessed a boom in research on economic growth theory, beginning with the work of Romer (1986) and Lucas (1988).

Selected Growth Studies Related to Africa

In the process of explaining Africa's growth problem, the World Bank has assembled a large database on many dimensions of Africa's development experience. Over the last decade, a growing number of development specialists have examined these data to better understand the statistical determinants of Africa's growth performance. The first study in this regard, by Easterly and Levine (1997), seeks explanations for the factors ascertaining the growth tragedy in Africa. The second study, by Radelet, Sachs and Lee (1997), analyzes the factors that have contributed to differences in growth rates between a sample of Asian and African countries.

The third, by Sachs and Warner (1997), attempts to measure the "sources of slow growth" in Africa. The fourth, by Block (1998), asks whether African countries "grow differently" from those in other regions. The fifth study, by Calamitsis, Basu and Ghura (1999), identifies empirically the main factors fostering adjustment and growth in Sub Saharan Africa (SSA). In the sixth study, Fosu (1999) explicitly notes that Africa's poor performance is the result of internal and external factors. Confining his attention to external factors, Fosu assembles evidence showing that Africa's exports have been determined exogenously and that exports have driven income growth. The final study by Easterly (1999) searches for reasons for the poor performance of developing countries in general. Easterly concludes that growth in developing countries had been systematically reduced by shocks that spilled over from developed countries.

These studies overlap in obvious ways. Taken together they help us identify many of the important factors that have affected economic growth in Africa. Since all of the studies rely on standard single equation growth regressions, their principal value is to highlight potentially fruitful associations between the explanatory variables and economic growth. Easterly and Levine derive a model of long term growth to analyze the variables that are directly and indirectly related to growth performance in Africa. They derive their basic equation from a model of long-term growth. The variables included in the equation are initial income, human capital, financial depth, black market exchange rate premium, central government surplus and several dummies relating to Africa's peculiarities.

Reviewing their results, Easterly and Levine concluded that the poor growth was strongly associated with (1) low schooling, (2) political instability, (3) under-developed financial systems, (4), distorted foreign exchange markets, as measured by the black market premium, (5), high government deficits, (6), low infrastructure, (7), ethnic fractionalization, and, (8), spillovers from neighbors that magnify (1) – (7).

The study by Radelet, Sachs and Lee (1997) examines cross-country differences in rates of growth between Africa and Asia. Their estimates highlight the relative importance for growth of efficient bureaucracy and institutions, good macroeconomic management, and strategies that enhance productivity. Using a growth accounting exercise for the period 1965 to 1990, the authors explain a significant portion of the difference in average annual growth rates under two headings, "policy variables" and "demography." The policy variables are (a) government savings

rate; (b) openness; and (c) institutions. The aggregate nature of their analysis confounds the effects of specific policy variables.

The statistical significance of the variable "institutions" points to the complex web of decisions, policies, and actions that enhance the efficiency of public bureaucracies, improve the competence of public sector administrators, promote effective implementation of policies and programs, maintain accountability, and enhance governance. The significance of the "government savings rate" is evidence of policies, decisions, and administrative actions that ensure governments conduct their affairs in ways that avoid (or overcome) distortions. The most common distortions that undermine growth in Africa are deficit financing, the rapid accumulation of domestic and foreign debt, ill-advised attempts to fix the exchange rate and interest rates, and interventions that hinder financial development. The variable "openness" represents policies and actions that enhance international competitiveness, promote sustained increases in total factor productivity, and encourage public and private investments that raise the level of output over time.

These results are suggestive. For example, using the estimated coefficients as a guide, there appears to be a direct link between economic growth (defined as sustained increases in real output per capita) and development (defined as generalized improvements in welfare). This is reflected in the significance of the demographic variable "life expectancy," an outcome consistent with a growing body of evidence suggesting that there is no trade-off between rapid growth and poverty reduction. On average, African countries have had exceedingly low growth rates, accompanied by increased poverty and welfare regression. By contrast, rapid growth in Asia has been accompanied by widespread poverty reduction and improving welfare.

In the work of Sachs and Warner (1997) the emphasis is on trade openness. They consider a sample of 74 countries in a cross-country regression for per capita growth between 1965 and 1990. They find that access to the sea, life expectancy, government savings, institutional quality and a growing population share of working age persons have a significant and positive influence on growth. Their results also show that resource endowments and a tropical climate impede growth. Sachs and Warner interpret their findings as evidence that growth in Africa is not different from growth elsewhere. The main reasons why African countries have grown slowly are that they are landlocked, predominantly tropical, have weak institutions, and have maintained counterproductive policies. The latter are evident in persistent budget deficits and commercial policies that close off African economies to international competition.

In the fourth study, Block (1998) inquires "Does Sub-Saharan Africa Grow Differently?" Seeking to move beyond analyses that treat SSA "primarily as a dummy variable in a single reduced-form growth regression", Block considers whether in Africa, the "mechanisms of economic growth operate differently". He does that using an "augmented reduced form" growth regression. The model is augmented by specifying separate equations for some explanatory variables in the growth regression. Block's growth regression includes initial per capita income, life expectancy at birth, a dummy for landlockedness, a political risk index, the growth rate of the net barter terms of trade, the Sachs-Warner index of openness, the overall budget deficit including grants, the difference between the population growth rate and the growth rate of the economically active population, real investment, and the growth rate of the population.

Block's results offer little that is new. Like Sachs and Warner, he concludes that countries in SSA do not grow differently from countries elsewhere. He does find, however, that

the factors influencing growth are weaker in SSA. He also finds that their effects have been undercut through inappropriate policies and institutional barriers. Block concludes that weak institutions and poor policies in SSA have been far more costly in terms of growth than in other regions.

Calamitsis, Basu and Ghura (1999) study begins with the optimistic view that some African countries are "on the move". They caution, however, that the social and economic situation in most African countries remains "fragile". For policy makers, the challenge is to focus on growth and poverty alleviation, and "integrate [Africa] fully into the world economy". The authors' goal is to determine the empirical impact of adjustment on economic growth (measured as the change in real per capita income). They use the results to suggest the types of changes needed to stimulate growth and reduce poverty.

Their growth regression includes initial income, population growth, ratios of private and government investment to GDP, index of human capital, dummy for sustained IMF programs, rate of inflation, standard deviation of inflation, central government budget deficit (excluding grants), change in real effective exchange rate, rate of export growth, percentage change in external terms of trade, index of political freedom, dummy for war, and series of country and time specific dummies. Expecting simultaneity bias due to endogenous regressors to be a problem, they run a number of tests.

Concluding that the tests show no such bias, they turn to their results. These show that private investment is a more robust determinant of growth than government investment. Human capital has a positive but not significant effect on income growth. And population growth has a major negative effect. The estimated coefficients of the budget deficit and real exchange rate are negative and that of export growth is positive. An interesting finding is that inflation has the correct (negative) sign but is not statistically significant. The authors also find that sustained implementation of IMF programs leads to an increase in per capita income growth.

Fosu (1999) study begins with the assertion that Africa's "uneven" growth performance has resulted from both internal and external factors. His analysis, however, focuses on the importance of external factors. In particular, he concentrates on questions related to "openness". Acknowledging that openness and the growth of exports are not the same, he nonetheless frames his analysis in terms of a growth accounting approach that defines income as a function of capital, labor, and exports. After some manipulation (logarithmic differentiation and several substitutions), Fosu derives the equation he estimates. It relates the growth of real income to the growth of labor, the ratio of investment to income, the growth of exports, and a term (the ratio of exports to non-exports) designed to measure the "externality" effects of trade. This equation is then estimated for a cross-section of African countries for the periods 1960-70 and 1970-80.

The results show that exports are positively related to the growth of income and that the coefficient is statistically significant. Fosu also concludes that external shocks, the real exchange rate, foreign aid, and debt were important determinants of growth. He suggests that debt had a threshold effect. Below a particular threshold of gross domestic investment to GDP, the level of debt raises the rate of growth; above the threshold, debt lowers the rate of growth. Fosu examines the endogeneity of exports and the direction of causation between growth and exports. He concludes that exports were exogenous and that causation ran from exports to income. There are

now several studies that reach the opposite conclusions (Rodrik 1998; Summers 1999; Frankel and Romer 1999).

Another study of Easterly (1999) titled "The Lost Decades: Explaining Developing Countries' Stagnation 1980-1998" begins with the observation that there was no change in the median per capita income in developing countries during the 1980s and 1990s. This contrasted with an increase of 2.5 percent recorded for the period 1960 to 1979. Easterly examines whether the loss of growth was the result of "(1) good policies that did not achieve desired results, (2) bad economic policies, or (3) some third factor like shocks?" Based on his evidence – cross-country regressions and comparison of turning points that relate events in the rich countries to those in the developing countries – he argues that the most likely explanation was point (3). The principal shock he finds was the "growth slowdown in the industrial world". This conclusion would resonate widely in African capitals. African leaders have persistently argued that their countries could not grow because of the impact of periodic shocks that originate outside Africa.

Englebert (2000) uses a very parsimonious empirical framework to consider per capita growth from 1960 to 1992 with a sample of 99 developing countries. His empirical model of growth includes only five significant variables: a lagged dependent variable, state legitimacy index, a developmental capacity index (modified to be orthogonal to state legitimacy), an East Asian dummy (which positively affects growth) and a tropical climate index. He provides a strong motivation for the relevance of this state legitimacy variable for explaining slow growth in African countries, but his econometric results are not very convincing due to the suspected omitted variable bias.

Englebert finds that the African dummy becomes an insignificant regressor when he includes a dummy for the historical legitimacy of the state. The state legitimacy variable is highly significant in his regressions, with a coefficient that is relatively stable around 0.02. Englebert shows that the significance of the African dummy is very sensitive to the inclusion of the state legitimacy variable: when this variable is included, the t-statistic on the coefficient of the African dummy turns insignificant. He also shows that legitimate states are more likely to have high scores on a range of indicators of institutional stability, good governance and prudent policymaking, including variables such as trade openness, the depth of the financial sectors, foreign indebtedness, enforceability of contracts, the risk of expropriation and civil liberties.

Most recent research on Africa's growth has been empirical. Generally, empirical estimation were based on augmented Solow growth model equation of the rate of output growth on the following variables, entering individually or in combination (i) a measure of the initial level of output and the initial level of technology to capture the impact of initial conditions; (ii) the [exogenous] rate of technological change to account for productivity changes; (iii) the savings rate to capture capital accumulation; (iv) the growth rate of the work force; (v) the rate of depreciation of capital; (vi) the share of capital in output; and, (vii) the rate of convergence to the steady-state (Barro and Sala-i-Martin, 1995). This specification is directly derived from a production function.

A number of empirical studies have found that the Solow growth model fails to explain Africa's economic growth. An "African dummy" has been found to be large and significant in cross-section studies, suggesting that Africa's growth responds to variables different from those explaining it elsewhere (Barro and Lee, 2010; Easterly and Levine, 1997). Other studies, as noted by Collier and Gunning (1999), eliminated the dummy "though to an extent by transferring the puzzle elsewhere". This is the case with Sachs and Warner (1997) for example, who do not

find a significant African dummy but instead find a significant "tropics dummy". Both specification and estimation techniques could explain the significance of the African dummy.

Most researchers have responded to the puzzle of the Africa dummy by re-specifying the growth model and adding variables thought to capture missing factors not explained by the textbook Solow model. First, some studies endogenize the savings variable by including in the model the policy variables influencing savings. These include the black market premium, the rate of inflation and the rate of the budget deficit. Even sociological variables such as ethnic fractionalization have been considered important in explaining the Africa's dummy (see Easterly and Levine, 1997). Sachs and Warner (1997) added geographical variables to the list and found a significant tropical dummy. More generally, some studies have also introduced political variables in growth models to explain better the growth process (see for instance Barro and Lee, 1993; Alesina et al., 1996; and Easterly and Levine, 1997). Hoeffler (2002) is among the few who responded to the debate over the African dummy from an econometric perspective. In her methodologically detailed study, Hoeffler found that the significance of the African dummy was due to estimation problems.

All the studies cited above used either cross section OLS or fixed effect panel approaches to estimate the growth model. However, it is simple to show that these methods are flawed when estimating dynamic panel data models. Hoeffler presents five models using five different estimation techniques. She finds that when the appropriate method of estimation is used, the African dummy is no more significant even when the model is restricted to the basic Solow model without adding any more variables. She therefore concludes that growth in Africa is explained by the same fundamental production function factors used in the Solow model.

Underlying the controversy is the complexity of the growth process. Most studies claiming to have explained growth account for just a small proportion of the variation in the rate of growth. This cannot be otherwise because growth has its country or regional idiosyncratic determinants. Whether these are so important that they invalidate the main pattern given by the basic variables of the Solow model is an empirical question. An important but rarely adopted approach to explaining growth, probably due to its high cost, remains the 'case study' approach. It is only through case study analysis that the predictions of cross-country models can be confronted with country 'realities' to determine their robustness.

THE EMPIRICAL MODEL

As revealed generally by the literature in the preceding section, modern real sector or optimal growth analysis recognizes the contributions of the real sector and its development policies, financial sector policies and performance, and, exogenous developments, to the growth process. Typical of the impulses from the real sector and related policies are the supply of labor input and the associated manpower development policy/program, and, process and product development policies. Governments contribute to manpower development policy and general human capital development through education and health programs and their related expenditures and also direct the growth process through economy-wide policies causing far reaching changes in many sectors. The quality of institutions falls into this latter category. From

the finance stable, short run growth drivers contributing to enhanced quality of investments through improved efficiency of available capital inputs, are supplied. In short, no key part of the economy is left out in the growth permutation.

Exogenous factors or events do set limits to growth; such limits could range in duration from short term to long term, the latter often engenders policies causing structural change as an escape from the penalizing factor or factors. The short term factors/events could include strife, religious/communal disturbances, drought causing famine and output drop and policy regime change with high initial adjustment costs. The long-term exogenous factors/events could take the form of population growth, technological change, deterioration in a country's terms of trade which may result from fall in the demand for export sales as a result of say, prolonged recession in the economies of foreign buyers or a permanent change in the taste of foreign buyers. While some of the short-term factors could be expected to adjust themselves hence calling for only short-term stabilization policies, the long-term factors would necessitate structural change policies [either to absorb – accommodation - or offset the exogenous or long-run shift factor] of the type mentioned above.

Aggregating these factors and re-arranging, the growth model to be estimated in this study is in the first instance specified as a panel data set of fifteen African countries such that:

1. $$GDP = f(K, L, INST, CNFL, PMP, RIR, INFL, DEBT, POPG, OPEN, TOT, TFPG, GDP_t)$$

(+) (+) (+) (-) (-) (+) (+/-) (+/-) (-) (+) (+) (+) (+)

Where, GDP is per capita real gross domestic product, K is capital input, L is labor input, INST is quality of institutions, CNFL is conflicts, PMERP is parallel market exchange rate premium, RIR is real interest rate, INFL is inflation rate, DEBT is overall debt exposure, POPG is population growth, TOT is external terms of trade, reflecting both domestic and foreign demands, TFPG is total factor productivity growth and GDP_t is trend nominal GDP. The signs underneath the variables denote a priori expectations.

In relation to the overview given in the introductory segment of this section, TFPG absorbs directly most of the governmental efforts in the real sector as it is directly influenced by the national system of innovations (NSI) encapsulating technological change (an exogenous and long-run factor) and other policies raising factor productivity in both the short and long runs. Such other policies include relative prices such as, exchange rate and interest rate which however are also designed to eliminate distortions in their relevant markets and thus could constitute independent sources of short run growth. This justified their separate inclusion in the model(In this study, exchange rate policy was to be proxy by the parallel market exchange rate premium (PMERP) which is generally believed to capture more aptly, the disequilibrium in the foreign exchange market. However, widespread data unavailability precluded its use; the real effective exchange rate was accordingly substituted. While, it could still be an effective indicator of policy distortion, it may not capture the direct effect of corruption discussed in the paper.).

TFPG would also be influenced (as per NSI) by schooling at all levels and other training and health programs hence, unlike other growth models, such factors are not viewed in this study

as independent sources of growth. Trend GDP is included as an exogenous variable capturing technological change in the sense of being the sole driver of consumers' surplus which reflects growth in the welfare sense (Ogun, 2012a).

An exogenous factor – strife, comprising of religious and communal disturbances, which is widespread in Africa, is represented by conflicts (CNFL).

In most studies of growth, corrupt practices are often emphasized/specified as short run determinant (see e.g. Mo, 2001; Mauro, 1995; 2004,). In the present study, the most significant impact of corruption is narrowed to that on TFPG where it exerts long run effect (see e.g. Ogun, 2012b). Accordingly, its direct growth effect was limited to the short run and reflected in the parallel market exchange rate premium serving as the incentive for 'round tripping' and other sharp practices in the financial and public sectors (see e.g. Ogun 2012c).

Debt (including fiscal deficit), inflation and openness represent the other policy factors (that is, quality of management) in the model, noting however, that, openness is a long run variable.

To some extent, both inflation and terms of trade would reflect the effect of weather condition with terms of trade also capturing the effect of taste.

In log expression, equation (1) becomes:

$$LogGDP = \alpha_0 + \alpha_1 LogK + \alpha_2 LogL + \alpha_3 LogINST + \alpha_4 LogCNFL + \alpha_5 LogPMERP + \alpha_6 LogRIR + \alpha_7 LogINFL + \alpha_8 LogDEBT + \alpha_9 LogPOPG + \alpha_{10} LogOPEN + \alpha_{11} LogTOT + \alpha_{12} LogTFPG + \alpha_{13} GDP_t + \mu_t \quad (2)$$

Where, the variables and the related partials are as earlier defined. An alternative specification in which a variable, $\frac{M2}{P}$, denoting real money balances (with expected positive effect on growth) is substituted for inflation appears as below (The alternated variables, that is, real money balances and inflation could not be contained in the same equation for obvious reason of mutlicollinearity – sustained inflation being a monetary phenomenon.).

$$LogGDP = \beta_0 + \beta_1 LogK + \beta_2 LogL + \beta_3 LogINST + \beta_4 LogCNFL + \beta_5 LogPMERP + \beta_6 LogRIR + \beta_7 \frac{M2}{P} + \beta_8 LogDEBT + \beta_9 LogPOPG + \beta_{10} LogOPEN + \beta_{11} LogTOT + \beta_{12} LogTFPG + \beta_{13} LogGDP_t + \varepsilon_t \quad (3)$$

The variables in the model and the relevant proxies are described below.

GDP = gross domestic product per capita;

K = capital stock defined as the sum of gross capital formation and personal consumption;

L = labor force defined as total annual employment;

INST = quality of institutions proxy by two indices, government effectiveness and regulatory quality. Both indices were obtained from the World Government Indicators (WGI) produced by Kaufmann, Kraay and Mastruzzi (2010). As

indicated by the authors, government effectiveness index reflects perceptions of the quality of public services, the quality of the civil service and the degree of its independence from political pressures, the quality of policy formulation and implementation, and the credibility of the government's commitment to such policies. Also, they described regulatory quality index as reflecting perceptions of the ability of the government to formulate and implement sound policies and regulations that permit and promote private sector development.

CNFL = social conflicts/strife proxy by an index of political stability and absence of violence and also obtained from WGI. According to the proponents, this index reflects perceptions of the likelihood that the government would be destabilized or overthrown by unconstitutional or violent means, including politically-motivated violence and terrorism.

PMERP = parallel market exchange rate premium proxy by real effective exchange rate;

RIR = real interest rate;

INFL = inflation defined as log difference of consumer price index;

M2/P = real money balances;

DEBT = the dollar value of the sum of total indebtedness – external and internal debt;

POPG = population growth;

OPEN = degree of openness conventionally represented as the ratio of the sum of exports and imports to gross domestic product (GDP);

TOT = terms of trade – relative price of exports and imports;

TFPG = total factor productivity growth defined as the change (percentage) in the sum of the ratio of gross national output (GNP) to total employment and the ratio of GNP to capital;

GDP_t = trend GDP generated as the fitted value of a regression of nominal GDP on time.

There is no particular yardstick employed in choosing the sample; the sample however reflects the different regions on the continent. The countries in the sample are: South Africa, Botswana, Mauritius, Kenya, Tanzania, Uganda, Democratic Republic of Congo, Gabon, Central African Republic, Nigeria, Ghana, Senegal, Egypt, Algeria and Tunisia.

THE RESULTS AND THEIR INTERPRETATIONS

As noted in the preceding section, both inflation and real money balances were interchanged in the estimation. Results were produced for the alternative specifications under a static model expressed in two forms: a log level specification and a log differenced dependent variable with log level explanatory variables. Also, three types of estimation results were produced: pooled (OLS), fixed effects and random effects. Under the static model, the Hausman statistics were significant suggesting a preference for the fixed effects approach. Nonetheless, the random effect estimates are retained for possible comparison. Besides the static model, dynamic panel estimations were also conducted. Still alternating real money balances and inflation, results were produced for differenced generalized method of moments (DIF-GMM) and system GMM (SYS-GMM).

The study covered the period 1996 to 2010. This scope imposed restrictions on the application of panel unit root and panel cointegration methodologies which would have signified the type of time-defined relationships that existed between the dependent and some independent variables in the model. However, as the trend of applications in the empirical literature suggests, these methodologies are of little significance when dealing with dynamic panel analysis. Nonetheless, an unorthodox way of inferring long-run relationship involving the static regression of theoretically identified steady-state variables on the dependent variable was explored. Unconventional, it at least gives an indication of the possibility of long-run relationships that might hold under the specifications.

Apart from data on political stability, government effectiveness and regulatory quality whose source has been reported, all the data employed in the study are from the World Development Indicators (2011) of the World Bank (The data set employed in the study is available from the author upon request).

The result of the static model corresponding to inflation in the list of explanatory variables is presented below.

Table 1. GDP Equation Considering Inflation

Static Model : GDP equation considering Inflation

Variable	Pooled (lngdp)	Fixed Effect (lngdp)	Random Effect (lngdp)		Pooled (D.lngdp)	Fixed Effect (D.lngdp)	Random Effect (D.lngdp)
Lnk	*0.3631***	*-0.044*	*0.3631718****		*0.0013*	*0.1047*	*0.0013*
Lnl	*-0.3461****	*0.7218****	*-0.3461252****		*-0.005*	*-0.451*	*-0.005*
Rir	*-3E-05*	*0.0013****	*-3E-05*		*0.0003*	*0.0001*	*0.0003*
Lnreer	*0.0323*	*-0.052*	*0.0323*		*0.0347*	*0.0457*	*0.0347*
Inflation	*-0.0075****	*-4E-04*	*-0.0075****		*-5E-04*	*-6E-04*	*-5E-04*
Lndebt	*-0.0679***	*-0.0353****	*-0.0679054***		*0.0116*	*0.0217*	*0.0116*
Lnpopg	*0.1120***	*0.0739***	*0.1120635***		*0.0223*	*-0.048*	*0.0223*
Tot	*4E-12*	*1.46E-11****	*4E-12*		*-6E-12*	*-1.16E-11***	*-6E-12*
Lntfpg	*0.8077****	*0.6478****	*0.8077****		*0.0019*	*0.063*	*0.0019*
Openness	*-8E-09*	*6.51E-9***	*-8E-09*		*4E-10*	*-2E-09*	*4E-10*
Polstab	*-0.068*	*-0.015*	*-0.068*		*-0.022*	*0.0037*	*-0.022*
Goveff	*-0.1725***	*0.0086*	*-0.1725957****		*-0.04*	*0.0436*	*-0.04*
Regqu	*0.1202***	*0.1000****	*0.1202***		*0.0644**	*0.0071*	*0.0644***
Lnfitted	*0.4987***	*-0.6136**	*0.4987***		*0.3585***	*0.7655*	*0.3585****
_cons	*-8.7305****	*-6.06***	*-8.7305****		*-3.0266***	*-2.346*	*-3.0266****
R-squared	*0.9984*	*0.0905*	*0.9984*		*0.4818*	*0.1494*	*0.4818*
F-stat	*1438.20****	*127.36****			*1.86**	*1.01*	
Hausman		*53.57****				*7.12*	

Note: here and in all tables ***, **, and, * denote significance at 1%, 5% and 10% respectively
Source: Computed

In table 1, variables such as real interest rate, terms of trade, openness, political stability, government effectiveness and regulatory quality were not entered in log due to their very small

values hence were entered in level. Under the OLS estimates, variables k, reer, infl, debt, tot, tfpg, polstab, regqu and fitted gdp entered with the correct sign and with the exception of polstab, all were significant at either 1 or 5 per cent. Notably, both inflation and debt generated adverse effect on gdp; rir, popg, a measure of institutional quality, goveff, l and openness were wrongly signed with the middle three highly significant. However, another measure of institutional quality, regqu, conformed to a priori expectation. The adjusted coefficient of multiple determination suggests that the explanatory variables accounted for over 99 per cent of the movements in the GDP.

With the fixed effects, variables k, reer, popg, fitted gdp bear the wrong sign with the last two significant at 5 percent. Rir and openness were now correctly signed and significant. However, only about 9 per cent of the variations in the GDP were explained by the independent variables. The results of the random effects are practically the same with the OLS estimates.

Comparatively, the case of the static model with differenced dependent variable while the independent variables remained at level was generally poor.

The static model estimates corresponding to real balances are presented in table 2.

Static Model : GDP equation considering m2/cpi						
Variable	**Pooled (lngdp)**	**Fixed Effect (lngdp)**	**Random Effect (lngdp)**	**Pooled (D.lngdp)**	**Fixed Effect (D.lngdp)**	**Random Effect (D.lngdp)**
lnk	*0.2893***	*-0.035*	*0.2893859****	*-3E-04*	*0.0944*	*-3E-04*
lnl	*-0.4184****	*0.4725226**	*-0.4184****	*-0.001*	*-0.208*	*-0.001*
rir	*0.0002*	*0.00115****	*0.0002*	*0.0003*	*0.0004*	*0.0003*
lnreer	*0.1280**	*-0.024*	*0.1280**	*0.041*	*0.0227*	*0.041*
m2cpi	*8.43E-12****	*2.36E-12***	*8.43E12****	*4E-14*	*-2E-12*	*4E-14*
lndebt	*0.0333**	*-0.018*	*0.0333**	*0.0122*	*0.0078*	*0.0122*
lnpopg	*-0.1460016****	*0.046*	*-0.1460****	*0.018*	*-0.02*	*0.018*
tot	*-4E-12*	*1.19E-11****	*-4E-12*	*-6E-12*	*-0.1E-11***	*-6E-12*
lntfpg	*0.6824697****	*0.6558965****	*0.6824697****	*0.0029*	*0.0317*	*0.0029*
openness	*-9.51****	*3E-09*	*-9.51E-11****	*2E-11*	*2E-10*	*2E-11*
polstab	*0.0141*	*-0.01*	*0.0141*	*-0.02*	*0.0068*	*-0.02*
goveff	*0.0361*	*-0.002*	*0.0361*	*-0.04*	*0.0356*	*-0.04*
regqu	*0.046*	*0.1056****	*0.046*	*0.0627**	*0.0193*	*0.0627**
lnfitted	*0.102*	*-0.361*	*0.102*	*0.3472***	*0.5664*	*0.3472***
_cons	*-3.5212***	*-4.7741***	*-3.5212****	*-3.0058***	*-3.674*	*-3.0058***
R-squared	*0.9993*	*0.1713*	*0.9995*	*0.4803*		
F-stat	*4472.40****	*155.57****		*1.85***		
Hausman		*128.65****			*9.12*	

Source: Computed

In Table 2, real balances joined the list of variables entered in level. With the OLS estimates, labor force, debt, terms of trade, openness and political stability (conflicts) were incorrectly signed. Contrarily, capital stock, real interest rate, real effective exchange rate, real money balances, population growth, total productivity growth, government effectiveness, regulatory quality and fitted gdp entered with the correct signs; capital stock, real money balances, population growth and total factor productivity growth were highly significant. The

fixed effects estimates appeared to follow the trend of the OLS with noticeable differences in the relative performances of real effective exchange rate, debt, terms of trade, openness and regulatory quality. The random effects estimates were not significantly different from the OLS. Again, the static model results corresponding to the differenced dependent variable and level independent variables were generally poor.

As indicated earlier, an attempt was made at assessing the pattern of long-run relations that may hold in the model by including only the theoretically defined long-run variables in a static model. The first set of results corresponding to level expression is presented below.

Table 3. Static Model I Considering Long-run Variables

Static Model: lngdp			
Variable	**Pooled**	**Fixed Effect**	**Random Effect**
lntfpg	*0.8389****	*0.8718****	*0.7606****
totr	*-1.08E-10****	*0.2E-11*	*-0.8E-13*
openness	*-3.11E-9****	*-9.31E-11**	*-1.55E-10***
lnfitted	*1.7714****	*0.4436****	*0.5437****
_cons	*-16.1152****	*-7.1389****	*-6.8751****
R-squared	*0.8916*	*0.7896*	*0.7955*
F-stat	*337.12****	*706.48****	
Hausman		*-27.31*	

Source: Computed

The results suggest that only two variables may play important long-run roles in the set of countries involved. These are, total factor productivity growth and fitted gdp. Terms of trade and openness are incorrectly signed even though significant hence, may not be credible long-run factors in the relevant countries.

The results of the differenced dependent variable are as follows.

Table 4. Static Model II Considering Long-run Variables

Static Model: D.lngdp			
D.lngdp	**Pooled**	**Fixed Effect**	Random Effect
lntfpg	*0.0054****	*0.0108*	*0.0047*
Tot	*-0.1E-11*	*0.2E-11*	*0.2E-12*
Openness	*0.4E-10*	*1.34E-9****	*2E-10*
lnfitted	*0.0492*	*-0.032*	*0.0197*
_cons	*-0.4133**	*0.0999*	*-0.193*
R-squared	*0.0755*	*0.0363*	*0.0649*
F-stat	*3.10***	*2.25**	
Hausman		*11.06****	

Source: Computed

With a highly significant Hausman statistic, the favored fixed effects estimates appear to suggest the possibility of openness being relevant in long-run growth consideration in the sample of countries.

With respect to the dynamic panels, the inflation and real balances interchange were also observed under the two GMM estimations, DIF-GMM (a) and SYS-GMM (b). The results corresponding to inflation are as presented below.

Table 5. Dynamic Model With Inflation

Variable	Inf[a](lngpd)	Inf[b](Ingdp)	inf[a](D.lngpd)	Inf[b](D.Ingdp)
Lngdp(-1)	*0.1937*	*0.4326****	*-0.531*	*-0.102*
lnk	*0.0148*	*0.2078****	*0.2238*	*0.0265*
lnl	*0.4097*	*-0.2391****	*-3.3592****	*-0.042*
rir	*0.0002*	*0.0014*	*0.0001*	*-0.002*
lnreer	*-0.044*	*0.0118*	*-0.007*	*0.0461*
inflation	*0.0002*	*0.0002*	*0.0012*	*0.0002*
lndebt	*-0.017*	*0.002*	*0.0095*	*0.0232*
lnpopg	*0.0544*	*0.0576*	*0.021*	*0.0863*
tot	*7E-12*	*-6E-12*	*-0.337E-10****	*-7E-12*
lntfpg	*0.5140****	*0.4078****	*0.5228***	*-0.045*
openness	*5E-09*	*-1E-09*	*-7E-09*	*2E-09*
polstab	*-0.01*	*0.0003*	*-0.009*	*-0.014*
goveff	*0.0451*	*-0.002*	*-0.055*	*0.0594*
regqu	*0.0055*	*-0.016*	*-0.01*	*-0.067*
lnfitted	*-0.275*	*0.3244*	*4.668907****	*0.1844*
_cons	*-5.075*	*-4.7144***	*6.004*	*-1.436*
Wald	*243.86****	*15505.81****	*19.41*	*9.71*
Sargan Test	*9.405856*	*10.0644*	*9.6471*	*11.0271*
Sargan Prob > chi2	*0.6679*	*0.9857*	*0.6469*	*0.9623*

Source: Computed

Under DIF-GMM, only four variables appeared to enter with the wrong sign viz: real effective exchange rate, inflation, population growth and fitted gdp. Of the remaining, total factor productivity growth was highly significant. The variable maintained this performance under SYS-GMM with initial GDP (lagged GDP), capital stock and labor force entering the significance list. With the differenced dependent variable, only total factor productivity growth and fitted gdp were credibly significant under DIF-GMM while no variable significance was recorded under SYS-GMM. Generally, the Sargan test statistics were insignificant suggesting some degree of appropriateness of the model especially as regards the choice of instruments.

With real balances, the results are as follow.

Table 6. Dynamic Model With Real Balances

Variable	m2cp[a](lngpd)	m2cpi[b](lngdp)	m2cpi[a](D.lngpd)	m2cpi[b](D.lngdp)
lngdp(-1)	*0.0561*	*0.3498328***	*-0.491*	*-0.251*
lnk	*0.0039*	*0.2244137****	*0.2042*	*0.0785*
lnl	*0.1576*	*-0.283 9584****	*-2.963223***	*-0.047*
rir	*0.0006*	*0.0015*	*-8E-05*	*-0.002*
lnreer	*-0.034*	*0.0226*	*-0.038*	*-0.014*
m2/cpi	*0.31E-11**	*0.2E-11*	*-0.1E-11*	*-0.4E-11*
lndebt	*-0.005*	*0.0084*	*0.0027*	*0.0007*
lnpopg	*0.0032*	*0.0115*	*0.0412*	*0.2373***
totr	*0.8E-11*	*-0.5E-11*	*-0.308E-10***	*-0.101E-10**
lntfpg	*0.7082****	*0.4608****	*0.432*	*-0.03*
openness	*0.1E-08*	*-0.2E-08*	*-0.4E-08*	*0.5E-08*
polstab	*-0.015*	*-0.007*	*-0.017*	*-0.009*
goveff	*0.0541*	*0.0143*	*-0.049*	*0.0146*
regqu	*0.0357*	*-0.006*	*-0.018*	*-0.111*
lnfitted	*0.0181*	*0.3116*	*4.2198***	*0.2448*
_cons	*-4.781*	*-4. 5381***	*4.9832*	*-2.588*
Wald	*132.24*	*16165.79****	*17.95*	*13.33*
Sargan Test	*8.8419*	*16165.79*	*9.4967*	*10.6490*
Sargan Prob > chi2	*0.7164*	*0.9854*	*0.6600*	*0.9692*

Source: Computed

Under DIF-GMM and with all variables in level, real balances and total factor productivity growth were significant. With SYS-GMM, total factor productivity growth, initial GDP, capital stock and labor force were significant. When the dependent variable was differenced one period, only the fitted gdp was credibly significant under DIF-GMM while no such equivalence was recorded under SYS-GMM.

An attempt was made to ascertain at different levels, the extent and direction of convergence in the models. First, unconditional convergence was tested with only the lag of the dependent variable in the equation. The outcome is reported below in table 7.

Table 7. Testing for Unconditional Convergence

```
      Source |       SS       df       MS              Number of obs =     210
-------------+------------------------------           F(  1,   208) =    4.81
       Model |   .00396965     1   .00396965           Prob > F      =  0.0294
    Residual |   .17159079   208  .000824956           R-squared     =  0.0226
-------------+------------------------------           Adj R-squared =  0.0179
       Total |   .17556044   209  .000840002           Root MSE      =  .02872

------------------------------------------------------------------------------
     D.lngdp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       lngdp |   .0035749   .0016297     2.19   0.029     .0003621    .0067878
       _cons |  -.0042303   .0111896    -0.38   0.706    -.0262897    .0178292
------------------------------------------------------------------------------
```

Source: Computed

The coefficient of lagged GDP is about 0.0036 with ($p<0.05$) suggesting that there is divergence among the African countries. High growing countries tend to grow more.

The case of conditional convergence was examined at various levels, first, with inflation and other explanatory variables minus fitted gdp and real balances. The outcome of the experiment is as follows.

Table 8. Conditional Convergence Considering Inflation without Fitted GDP

```
      Source |       SS       df       MS              Number of obs =      43
-------------+------------------------------           F( 14,    28) =    1.16
       Model |  .013641373    14  .000974384           Prob > F      =  0.3541
    Residual |  .023482704    28  .000838668           R-squared     =  0.3675
-------------+------------------------------           Adj R-squared =  0.0512
       Total |  .037124077    42  .000883907           Root MSE      =  .02896

------------------------------------------------------------------------------
     D.lngdp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       lngdp |   .0819009   .0929553     0.88   0.386    -.1085094    .2723113
         lnk |  -.0231872   .0350922    -0.66   0.514    -.0950703    .0486959
         lnl |   .0311281   .0462769     0.67   0.507    -.0636657     .125922
         rir |   .0000729   .0006139     0.12   0.906    -.0011847    .0013305
      lnreer |   .0270976   .0712945     0.38   0.707    -.1189426    .1731379
   inflation |   .0009257   .0020656     0.45   0.658    -.0033056    .0051569
      lndebt |  -.0013652   .0149034    -0.09   0.928    -.0318934     .029163
      lnpopg |   .0072716   .0292881     0.25   0.806    -.0527225    .0672656
        totr |  -1.52e-12   3.75e-12    -0.40   0.689    -9.19e-12    6.16e-12
      lntfpg |  -.0598544   .0783267    -0.76   0.451    -.2202993    .1005906
    openness |   1.46e-09   3.44e-09     0.43   0.674    -5.59e-09    8.52e-09
     polstab |  -.0052472   .0314227    -0.17   0.869    -.0696137    .0591192
      goveff |  -.0232771   .0473818    -0.49   0.627    -.1203344    .0737802
       regqu |   .0372632   .0340522     1.09   0.283    -.0324896    .1070159
       _cons |   .1035367   .8195697     0.13   0.900    -1.575276    1.782349
------------------------------------------------------------------------------
```

Source: Computed

The coefficient of lagged GDP is positive and insignificant suggesting divergence.

With the inclusion of fitted gdp, the outcome is shown below in Table 9.

Table 9. Conditional Convergence Considering Inflation with Fitted GDP

```
      Source |       SS       df       MS              Number of obs =      43
-------------+------------------------------           F( 15,    27) =    1.67
       Model |  .017887777    15  .001192518           Prob > F      =  0.1185
    Residual |    .0192363    27  .000712456           R-squared     =  0.4818
-------------+------------------------------           Adj R-squared =  0.1940
       Total |  .037124077    42  .000883907           Root MSE      =  .02669

------------------------------------------------------------------------------
     D.lngdp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       lngdp |  -.0055571   .0928637    -0.06   0.953    -.1960976    .1849834
         lnk |   .0032752    .034112     0.10   0.924    -.0667168    .0732672
         lnl |  -.0062124   .0453123    -0.14   0.892    -.0991855    .0867607
         rir |   .0003379   .0005762     0.59   0.562    -.0008443    .0015201
      lnreer |   .0346559   .0657842     0.53   0.603    -.1003221    .1696339
   inflation |  -.0005774    .002001    -0.29   0.775     -.004683    .0035282
      lndebt |   .0113645   .0146926     0.77   0.446    -.0187822    .0415112
      lnpopg |   .0228226   .0277359     0.82   0.418    -.0340867    .0797319
        totr |  -5.78e-12   3.87e-12    -1.49   0.147    -1.37e-11    2.16e-12
      lntfpg |   .0063449   .0771172     0.08   0.935    -.1518865    .1645762
    openness |   3.32e-10   3.21e-09     0.10   0.918    -6.25e-09    6.91e-09
     polstab |  -.0224599   .0298077    -0.75   0.458    -.0836203    .0387005
      goveff |   -.041263   .0442883    -0.93   0.360    -.1321351    .0496091
       regqu |   .0648355   .0333557     1.94   0.062    -.0036046    .1332757
    lnfitted |    .361983   .1482711     2.44   0.021     .0577559    .6662101
       _cons |  -3.083619   1.508277    -2.04   0.051    -6.178348    .0111092
------------------------------------------------------------------------------
```

Source: Computed.

The coefficient on lagged GDP is negative but insignificant suggesting convergence.

Under the real balances, the estimates without the inclusion of the fitted gdp are reported in the table below.

Table 10. Conditional Convergence with M2/CPI but without Fitted GDP

```
      Source |       SS       df       MS              Number of obs =      43
-------------+------------------------------           F( 14,    28) =    1.15
       Model |  .013556478    14   .00096832           Prob > F      =  0.3620
    Residual |  .023567599    28    .0008417           R-squared     =  0.3652
-------------+------------------------------           Adj R-squared =  0.0478
       Total |  .037124077    42  .000883907           Root MSE      =  .02901

------------------------------------------------------------------------------
     D.lngdp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       lngdp |   .0121392   .2173026     0.06   0.956    -.4329851    .4572634
         lnk |  -.0004366   .0641482    -0.01   0.995    -.1318381     .130965
         lnl |  -.0058439   .0977292    -0.06   0.953     -.206033    .1943451
         rir |   .0000873   .0006173     0.14   0.889    -.0011771    .0013517
      lnreer |   .0222634   .0698735     0.32   0.752     -.120866    .1653929
       m2cpi |   6.15e-13   1.95e-12     0.32   0.755    -3.38e-12    4.61e-12
      lndebt |   .0008097   .0190099     0.04   0.966    -.0381303    .0397498
      lnpopg |   .0032357   .0411593     0.08   0.938    -.0810753    .0875468
        totr |  -1.22e-12   3.65e-12    -0.34   0.740    -8.70e-12    6.26e-12
      lntfpg |  -.0164345   .1492374    -0.11   0.913    -.3221336    .2892646
    openness |   1.57e-09   3.56e-09     0.44   0.662    -5.72e-09    8.86e-09
     polstab |   -.007193   .0308788    -0.23   0.817    -.0704453    .0560592
      goveff |  -.0147617   .0545215    -0.27   0.789    -.1264439    .0969204
       regqu |   .0361485   .0348496     1.04   0.308    -.0352377    .1075346
       _cons |   .0790603   .8829818     0.09   0.929    -1.729646    1.887767
------------------------------------------------------------------------------
```

Source: Computed

The result clearly suggests divergence among the countries in the sample.

Finally, the exercise involving real balances and fitted gdp yields the following outcome.

Table 11. Conditional Convergence with M2/CPI and Fitted GDP

```
      Source |       SS       df       MS              Number of obs =      43
-------------+------------------------------           F( 15,    27) =    1.66
       Model |  .017832305    15   .00118882           Prob > F      =  0.1211
    Residual |  .019291772    27   .00071451           R-squared     =  0.4803
-------------+------------------------------           Adj R-squared =  0.1916
       Total |  .037124077    42  .000883907           Root MSE      =  .02673

------------------------------------------------------------------------------
     D.lngdp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       lngdp |  -.0112334     .20044    -0.06   0.956    -.4225023    .4000356
         lnk |   .0029432   .0591192     0.05   0.961    -.1183593    .1242458
         lnl |  -.0056572    .090043    -0.06   0.950    -.1904101    .1790957
         rir |   .0003322   .0005775     0.58   0.570    -.0008527     .001517
      lnreer |    .041981   .0648807     0.65   0.523    -.0911432    .1751052
       m2cpi |   1.33e-13   1.81e-12     0.07   0.942    -3.58e-12    3.84e-12
      lndebt |   .0126712   .0181736     0.70   0.492    -.0246179    .0499604
      lnpopg |   .0166897   .0383189     0.44   0.667    -.0619343    .0953137
        totr |  -5.89e-12   3.87e-12    -1.52   0.140    -1.38e-11    2.05e-12
      lntfpg |   .0105124   .1379407     0.08   0.940    -.2725186    .2935434
```

openness	-9.39e-11	3.35e-09	-0.03	0.978	-6.97e-09	6.78e-09
polstab	-.0193043	.0288778	-0.67	0.509	-.0785566	.039948
goveff	-.0387616	.0511825	-0.76	0.455	-.1437795	.0662563
regqu	.0624867	.0338658	1.85	0.076	-.0070002	.1319735
lnfitted	.3476795	.1421259	2.45	0.021	.0560612	.6392978
_cons	-3.040245	1.512541	-2.01	0.055	-6.143722	.0632332

Source: Computed

The result clearly suggests convergence.

CONCLUDING OBSERVATION

The results of the analyses in this paper generally supported the established view in the literature on the importance of capital abundance, labor supply, institutions, factor productivity and real balances in the growth process of African countries. Inflation, policy distortions, conflicts and debt (total) were negative influences. The test on institutions accepted the alternative hypothesis of reduced institutional weaknesses improving economic growth.

The performance of total productivity growth was unexpected and could in the first instance be interpreted as suggesting a departure from the standard view of a declining productivity growth calling forth explanation(s) perhaps, in the manner of its computation in this study. An eclectic interpretation which is consistent with the established view in the literature would underscore its unparalleled importance in the growth process as underwritten by its remarkable performance in this study. This therefore throws a challenge at African governments on the state of their national system of innovations. Clearly, an essential ingredient to achieving continuous improvement in productivity growth is the need to raise further, the promotion of an enhanced national system of innovations.

The generally poor performance of terms of trade in the results may be a reflection of the reality of the composition of African trade being mostly primary exports and finished goods import with well-known adverse price movements. Thus, an accelerated program of transition from primary to secondary goods production and export would be growth beneficial.

The evidence on convergence was mixed; unconditional convergence proposition was not supported but conditional convergence was obtained only with the presence of fitted gdp signifying the importance of rapid technological progress in African countries' desire to catch up with the more advanced economies.

REFERENCES

Alesina, A. and Perotti, R. 1996a. Income Distribution, Political Instability, and Investment. *European Economic Review* 40(6), 1203-1228

Alesina, A., Ozler, S. Roubini, N. and Swagel, P. 1996b. Political Instability and Economic Growth. *Journal of Economic Growth* 1 (June), 189-211.

Barro, Robert and Xavier Sala-i-Martin. 1995. *Economic Growth*. New York: McGraw-Hill.

Barro, R. J. and Lee, J.-W. 2010, A new data set of educational attainment in the world, Working Paper 15902, *National Bureau of Economic Research*. URL**:** http://www.nber.org/papers/w15902

Berthelemy, J.-C. and Söderling, L. 2001, The role of the capital accumulation, adjustment and structural change for economic take-off : Empirical evidence from African growth episodes, *World Development* 29(2), 323–343.

Block, S.A. 1998. Does Sub-Saharan Africa Grow Differently? Draft, Fletcher School of Law and Diplomacy, November.

Calamitsis E. A., Basu, A. and Ghura, D. 1999. "Adjustment and Growth in Sub-Saharan Africa" IMF *Working Paper* WP/99/31, April.

Cass, D., 1965. Optimum growth in an aggregative model of capital accumulation. *Review of Economic Studies,* 32(3): 233-240.

Collier, P. and Gunning, J. W. 1999, Why has Africa grown so slowly?, *Journal of Economic Perspectives* 13(3), 3–22.

Domar, E.D., 1946. Capital expansion rate of growth and employment. *Econometrica*, 14(2): 137-147.

Easterly, W. and Levine, R. 1997, Africa's growth tragedy: Policies and ethnic divisions, *The Quarterly Journal of Economics* 112(4), 1203–1250.

Easterly, W. 1999. The Lost Decades: Explaining Developing Countries' Stagnation 1980-1998. Draft Policy Research Group, *World Bank,* November.

Englebert, P. 2000. *State Legitimacy and Development in Africa*. Boulder: Lynne Reinner.

Fosu, A.K. 1999. The International Dimension of Growth in Sub-Saharan Africa. Paper presented at the AERC/Harvard Conference on "Explaining African Economic Growth Performance" March 26-27, 1999, Harvard University.

Fosu, A. K. 2002. Political instability and economic growth: Implications of coup events in Sub-Saharan Africa, *American Journal of Economics and Sociology* 61(1), 329–348.

Frankel, J. A. and D. Romer. 1999. Does Trade Cause Growth?. *The American Economic Review* 89(3), 379-399.

Harrod, R.F., 1939. An essay in dynamic theory. *Economic Journal*, 49: 14-33.

Hoeffler, A. E.: 2002. The augmented Solow model and the African growth debate, *Oxford Bulletin of Economics and Statistics* 64(2), 135–158.

Knight, F., 1944. Diminishing returns from investment. *Journal of Political Economy,* 52(1), 26-47.

Koopmans, T.C., 1965. On the concept of optimal economic growth. Yale University, New Haven, England.

Kaufmann, D. Kraay, A. and Mastruzzi, M. 2010. The Worldwide Governance Indicators : A Summary of Methodology, Data and Analytical Issues. World Bank Policy Research Working Paper No. 5430 http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1682130

Lucas, R.E. Jr., 1988. On the mechanics of economic development. *Journal of Monetary Economics*, 22(1), 3-42.

Malthus, T.R., 1798. An Essay on the Principle of Population. Electronic Scholarly Publishing , London.

Mauro, P., 1995. Corruption and Growth. *The Quarterly Journal of Economics*, 110(3), 681-712. doi:10.2307/2946696

Ndulu, B. J. and O'Connell, S. A.: 1999, Governance and growth in Sub-Saharan Africa, *Journal of Economic Perspectives* 13(3), 41–66.

Ndulu, B. J. and O'Connell, S. A. 2009, Policy plus: African growth performance 1960-2000, in B. J. Ndulu, S. A. O'Connell, R. H. Bates, P. Collier and C. C. Soludo (eds), *The Political Economy of Economic Growth in Africa*, 1960-2000, Vol. 1 of Cambridge Books, Cambridge University Press.

Ogun, O.D. 2012a, Money, Finance and Growth: A Critical Review, Department of Economics, University of Ibadan.

Ogun, O. D. 2012b, Exchange Rate Determination in Developing Economies, *Modern Economy* 3(5), 518-521.

Ogun, O.D. 2012c, Corruption and Long-run Growth: An Analytical Framework, Unpublished.

Radelet, S.C., Sachs, J.D. and J-W Lee.1997. Economic Growth in Asia. Draft, Harvard Institute for International Development, Cambridge

Ramsey, F., 1928. A Mathematical Theory of Saving. *Economic Journal*, 38(152), 543-559.

Ricardo, D. and Eck, G.1817. The Principles of Political Economy and Taxation. Prometheus Books, London.

Rodrik, D.1998. Trade Policy and Economic Performance in Sub-Saharan Africa.
NBER Working Paper No. 6562, May.

Romer, P.M., 1986. Increasing returns and long-run growth. *Journal of Political Economy*, 94(5), 1002-1037.

Sachs, J. D. and Warner, A. M. 1997, Sources of slow growth in African economies, *Journal of African Economies* 6(3), 335–76.

Schumpeter, J., 1934. The Theory of Economic Development: An Inquiry into profits, capital, credit, interest and the Business cycle. Harvard University Press, Harvard.

Smith, A., 1991. Wealth of Nations. Prometheus Books, New York, ISBN-10: 0879757051.

Solow, R., 1956. A contribution to the theory of economic growth. *Quarterly Journal of Economics,* 70(1), 65-94.

Summers, L.H., 1999. Reflections on Managing Global Integration. *Journal of Economic Perspectives,* 13(2), 3-18.

Swan, T.W., 1956. Economic growth and capital accumulation. *Economic Record*, 32(2), 344-361.

Young, A ., 1928. Increasing returns and Economic progress. *Economic Journal*: 38(152), 527-542.

EXCHANGE RATES AND TOURISM: EVIDENCE FROM THE ISLAND OF GUAM

Maria Claret M. Ruane, University of Guam

ABSTRACT

Guam is a U.S. territory in the Western Pacific region. It is a small island economy that, like many island economies around the world, lacks diversification and mostly relies on a few economic activities, especially tourism. Worse yet, Guam's tourist markets also lack diversification, with approximately 70% accounted for by tourists from Japan. With the significantly stronger U.S. dollar (USD) and weaker Japanese yen (JPY) since September 2012, the cost to Japanese tourists of visiting Guam had increased by 33%.

Given Guam economy's heavy reliance on Japanese tourism, this study aims to use available time series data and Ordinary Least Squares regression models to quantify the effect of the significantly stronger USD/weaker JPY in the past year on the number of Japanese tourists visiting Guam. The results of this study will be useful in formulating economic policies in Guam and also in other economies that are similar to Guam for their use of the USD as their local currency or as a peg to their local currencies as well as their tourist-oriented economies that cater to Japanese tourists.

INTRODUCTION

Guam is a U.S. territory in the Western Pacific region. It is an island economy that is small both in terms of its economic size (its latest real GDP at $4 billion in 2005 prices) and in terms of its population (160,000 residents according to the 2010 U.S. Census data). Like many island economies around the world, Guam's economy lacks diversification and mostly relies on a few economic activities, one of them being tourism. In 2012, Guam was destination to 1.3 million tourists, with approximately 70% of these tourists visiting from Japan.

In the past year, Japan's central bank, i.e., the Bank of Japan (BOJ), has pursued a policy of increasing money supply in order to boost Japan's economy, which has been sluggish for 15-20 years. This policy is designed to fight the deflationary tendencies of Japan's economy by raising the inflation rate to its target of 2% per year. As a result of this policy, the U.S. dollar (USD) has strengthened and the Japanese yen (JPY) has weakened significantly from 1USD = 77.61 on September 28, 2012 to 1USD to 103.18 JPY on May 23, 2013. This represented a 33% stronger USD/weaker JPY. For Japanese visitors who make purchases in USD, including those who visit Guam and other locations that use the USD as their local currency, the JPY cost had

just increased 33%, even if the USD prices have not change. Since then, the Japanese yen has fluctuated around 100 JPY to 1USD, the exchange rate that the BOJ and many Japan economy experts believe is the exchange rate that will boost domestic spending in Japan's economy sufficiently to yield a 2% inflation rate.

Given Guam's heavy reliance on Japanese visitors, this study aims to quantify and analyze the effect of the significantly stronger USD/weaker JPY in the past year on the number of Japanese tourists visiting Guam. The study is organized as follows. Section II presents an overview of Guam's economy, which highlights its lack of economic diversification. It also provides details on Guam's tourism sector, to which Japanese tourists contribute a large share. Section III starts off more generally by presenting the theoretical background on the relationship between exchange rates and tourism and then proceeds to narrow down the focus to changes in the exchange rates between the USD and the JPY in the past five year but, more importantly, in the past year. Section IV reviews the literature on the relationship between exchange rates and tourism, which confirms that many studies used tourist arrival to a destination economy as the dependent variable, and considered the effects of independent/explanatory variables such as exchange rates (which is the variable of interest in this study), tourists' income and others variables on tourist arrival data. The review of the literature shows that no previous study of this type for Guam exists and that this study fills this gap. Section V constructs an empirical model for analyzing the effect of the exchange rate between the USD and the JPY on Japanese tourist arrival in Guam and discusses the results of using monthly data from October 2003 to July 2013 in Ordinary Least Squares regression models. Section VI concludes the study and discusses policy recommendations.

GUAM'S ECONOMY AND TOURISM

Guam is an island economy that is small both in terms of its economic size (its latest real GDP at $4 billion in 2005 prices, U.S. Department of Commerce, Bureau of Economic Analysis, 2012, September 24) and in terms of its population (160,000 residents according to the 2010 U.S. Census data). These figures suggest Guam's annual per capita real income of USD25,000 in 2005 prices.

Like many island economies around the world, Guam's economy lacks diversification and mostly relies on a few economic activities that serve primarily three groups of customers: local residents, U.S. Federal government (including military) personnel and their families, and tourists.

Local Residents

Local residents provide strong support for retail trade and many different service industries in Guam, including health, education, financial, legal, etc. Another advantage of this class of customers is their contribution to the overall economy tends to be more stable and less

vulnerable to external shocks that affect the other two economic activities, U.S. Federal Government, including Military, which depends on congressional decisions and budgetary resources from Washington, D. C., and also affected by U.S. economic, political and military allies around the world; tourism in Guam depends on economic and other factors (including natural disasters) that affect countries and economies from where tourists originate.

U.S. Federal Government including Military Personnel and Family Members

As a U.S. territory, Guam benefits from receiving funding from the U.S. Federal Government for a wide array of activities, including the military presence on the island. The U.S. Federal Government contributed 41% of Guam's approximately USD4 billion real GDP in 2010 (U.S. Department of Commerce-Bureau of Economic Analysis, 2012 September 24) and accounts for 6.7% of 60,220 total employment in Guam in June 2013 (Guam Department of Labor-Bureau of Labor Statistics).

Tourists

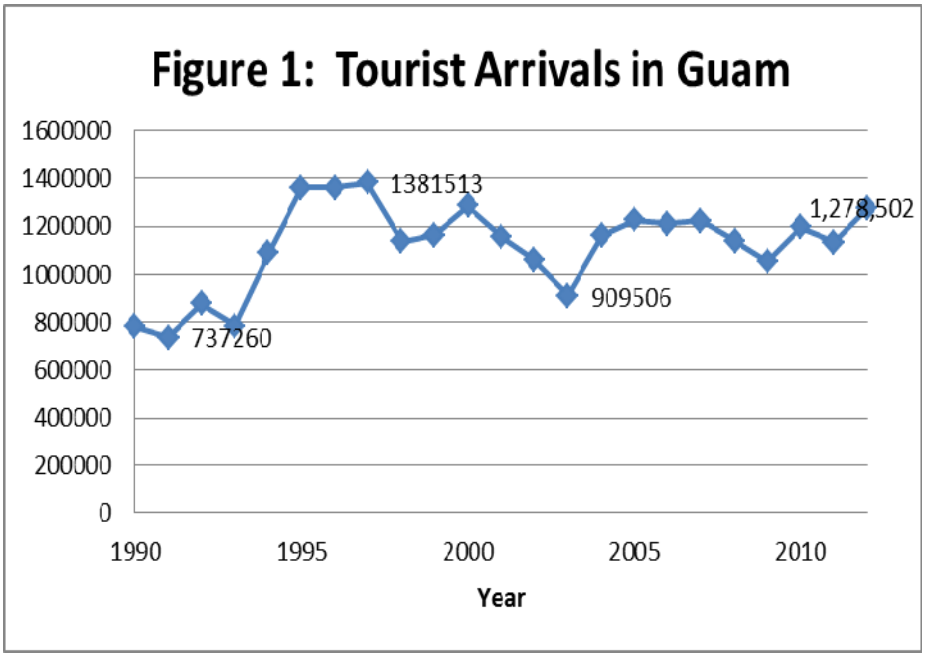


Figure 1 shows the number of annual visitors to Guam between 1990 and 2012. First is to note the overall volatility of the data, which highlights the fact that tourism in Guam and many economies is subjected to many external factors. Second is that Guam has been attracting at least one million visitors per year since 1994, with the exception of 2003. Third is that the peak, i.e., the largest number of visitors to Guam, occurred in 1997, the year of the Asian Crisis, which

explains the sharp decline that followed. Since 2004, tourist arrivals have fluctuated around 1.2 million visitors per year.

Tourist Markets

For years, the majority of visitors to Guam come from Japan, although this share has decreased from as high as 85-90% decades ago. The most recent data for the current fiscal year-to-date (October 2012 to July 2013) show that Guam welcome 1.087 million visitors who arrived by air (a small number, i.e., 7,029 visitors, arrived by sea). Of those who arrived by air, 68.39% were from Japan, 17.45% from Korea, 3.75% from the U.S. Mainland, i.e., the 48 U.S. states, 3.57% from Taiwan, with the remainder accounted for by smaller shares from other origin countries and is reported in Table 1. The share of Japanese visitors is the lowest in decades, or even compared to the last 5 years where it would be as high as 74%.

Table 1 also shows that visitors to Guam who arrived by air increase 6.5% compared to the same period a year ago. In terms of growth of individual origin countries, Table 1 shows that fastest growing tourist segments to be Russian visitors, who have enjoyed eligibility to the Visa Waiver Program to Guam and the U.S. since January 2012. Other fast growing groups to visit Guam are Korean and European visitors (each market grew 41.4% more this year than last year), Chinese visitors from Mainland China (15.9% higher than last year), and from Hong Kong (8.1% higher than last year).

Tourist Spending

One of the economic benefits to the destination economy (Guam, in this case) of tourism is the amount that tourists spend during their visit. Note that this is only part of the total spending that tourists contribute to the destination economy but represents the most direct benefit of tourism to the destination economy. The reason for this is that tourists also have prepaid expenditures, especially for accommodations and meals, which are not factored into the calculation below because of the complexity of calculating how much of the prepaid expenditures ultimately ends up in the local economy, especially when hotels providing the accommodations are foreign-owned and repatriate their revenue and/or profit to their home country.

Table 1: Guam Tourism Data			
	Oct 2012 - Jul 2013	**Share of Total Arrivals**	**% increase from a year ago**
TOURISTS IN GUAM BY MODE OF ENTRY			
Air Arrivals	1,087,211	99.36%	6.50%
Sea Arrivals	7,029	0.64%	44.20%
TOTAL TOURIST ARRIVALS	1,094,240	100.00%	6.70%
TOURISTS IN GUAM BY ORIGIN COUNTRY	**Oct2012-Jul2013**	**Share of Air Arrivals**	**% increase from a year ago**
JAPAN	743,582	68.39%	1.90%
KOREA	189,707	17.45%	41.40%
CHINA	8,540	0.79%	15.90%
HONG KONG	7,742	0.71%	8.10%
TAIWAN	38,799	3.57%	-6.90%
U.S. MAINLAND	40,756	3.75%	-7.00%
HAWAII	7,815	0.72%	-18.50%
CNMI	12,823	1.18%	-11.60%
PALAU	2,539	0.23%	-18.00%
FSM	8,242	0.76%	-1.40%
RMI	750	0.07%	-12.90%
PHILIPPINES	9,060	0.83%	3.40%
AUSTRALIA	2,786	0.26%	-18.70%
CANADA	784	0.07%	13.60%
EUROPE	1,836	0.17%	41.40%
THAILAND	310	0.03%	-7.70%
VIETNAM	72	0.01%	-18.20%
RUSSIA	5,530	0.51%	145.60%
OTHER/UNKNOWN	5538	0.51%	52.90%
TOTAL TOURIST ARRIVALS BY AIR	1,087,211	100.00%	
Source: Guam Visitors Bureau (various issues). Visitor Arrivals Statistics. Retrieved from http://www.visitguam.org			

Table 2 shows an estimate for this amount for Guam to be USD574.28 million for the current fiscal year, which accounts for spending of 90.63% of the total number of tourists that is expected to visit Guam this current fiscal year. Scaled to 100%, the amount comes out to be USD633.65 million of total tourist expenditure in fiscal year 2013. Using the spending multiplier of 1.3 (Ruane, 2011, December), which means every dollar spent on Guam multiplies demand and income in the local economy and ultimately generates an additional 30 cents of

spending and income. Therefore, the USD633.65 million of tourist expenditures for the fiscal year 2013 is expected to increase Guam's Nominal Gross Domestic Product (GDP), which was last estimated in 2010 at USD4.577 billion, by USD823.75 million or 18%.

Table 2: Tourist Expenditure in Guam

	Share of total number of tourists (from Table 1)	**Average days of visit per tourist***	**In-Guam expenditure per tourist***	**Estimated number of tourists for FY2013****	**Tourist expenditure in FY2013****
Japan	68.39%	2.85	$ 496.38	918815	$ 456,081,378.28
Korea	17.45%	3.31	$ 362.87	233057	$ 84,569,393.59
Hong Kong	0.71%	2.31	$ 198.71	9076	$ 1,803,491.96
Taiwan	3.57%	3.35	$ 424.13	46411	$ 19,684,297.43
Russia	0.51%	14.33	$ 1,687.39	7196	$ 12,142,458.44
	90.63%				**$ 574,281,019.70**

Sources:
*Guam Visitors Bureau (2013, April-June). Hong KongVisitor Tracker Exit Profile, prepared by Qmark Research. Retrieved from http://www.visitguam.org
*Guam Visitors Bureau (2013, June). Japan Visitor Tracker Exit Profile, prepared by Qmark Research. Retrieved from http://www.visitguam.org
*Guam Visitors Bureau (2013, July). Korea Visitor Tracker Exit Profile, prepared by Qmark Research. Retrieved from http://www.visitguam.org
*Guam Visitors Bureau (2013, January-March). Russia Visitor Tracker Exit Profile, prepared by Qmark Research. Retrieved from http://www.visitguam.org
*Guam Visitors Bureau (2013, April-June). Taiwan Visitor Tracker Exit Profile, prepared by Qmark Research. Retrieved from http://www.visitguam.org

Notes: ** author's calculation

As the economy expands, more jobs are created. Keeping the estimates to the year 2010 in the absence of more recent data, data show that Guam's USD4.577 billion economy created 62,600 jobs, or 1 job for every USD73,115 worth of economic activity. Based on the estimated increase in Guam's GDP resulting for tourist expenditures for fiscal year 2013, it is expected to have created 8,666 jobs, accounting for 13-14% of jobs in the Guam economy.

In addition to jobs created by tourism, additional taxes are collected by the government of the destination economy, which then finance a wide array of economic and social programs for the local residents. The two most obvious taxes earned by the local government from the additional GDP resulting from tourist expenditures for fiscal year 2013 are Gross Receipts Tax

(GRT) and Hotel Occupancy Tax (HOT). On Guam, the GRT rate is 4% of the total amount spent for most goods and services (including hotel services) and already reflected in the price paid by tourists and other consumers, and the HOT rate is 11% of the amount spent on hotel accommodations. Based on the additional GDP of USD823.75 million noted above, 4% of this is approximately USD33 million worth of GRT for fiscal year 2013. For HOT, the estimate is given by the Guam Visitors Bureau as USD 20.34 million for the period of October 2012 to July 2013, with two months left in the current fiscal year, this amount is estimated to be approximately USD24 million for the entire fiscal year 2013. Note that this tax calculation does not include other taxes, which would include additional personal and corporate income taxes imposed on the increased economic activity and incomes resulting from tourist expenditures estimated above.

All these benefits are summarized in Table 3.

Table 3: Economic Benefits from Tourism in Guam
Based on USD633.65 million of tourist expenditures in fiscal year 2013

Type of Benefit	Estimated amount
Direct, indirect and induced spending and income	USD823.75 million (18% of Guam's GDP)
Taxes due to local government (GRT and HOT)	USD 57 million
Jobs created	8,666 jobs (13-14% of total jobs)
Note: author's calculation	

EXCHANGE RATES AND TOURISM

Theoretical Background

The nominal exchange rate is defined to be the number of local currency used to buy/exchange for a foreign currency. Since this study involves only two currencies (USD and JPY), this measure of exchange rate is appropriate to use. This measure also works well when the inflation rates in the two countries are low, which is the case for the U.S. and Japan, so that the differential inflation rate, when it exists, is minimal; otherwise, the more appropriate measure of exchange rate would be the real exchange rate. If more countries and their currencies are involved, most studies use a weighted average of the changes in the real exchange rates among the currencies involved (Crouch, 1993, page 48).

One sees that the nominal exchange rate represents a bilateral (two-sided) relationship: For Japanese tourists who spend USD during their visit to Guam, their local currency is JPY and foreign currency is USD. To Guam residents, their local currency is USD and foreign currency is JPY. When one currency (in this case, USD) strengthens, the other currency (JPY) weakens, which means one requires more JPY now than before to buy the same 1USD or to pay for products priced in USD, even if the USD price has not changed. For example, an item that is priced USD100 would have cost JPY7,600 in September 2012 but would cost JPY10,000 now

that the exchange rate is around JPY100 to 1USD. As is illustrated in this example, a stronger USD translates to a weaker JPY, which means that the cost to a Japanese visitor to Guam has increased. With this higher cost, it is hypothesized to affect Japanese tourism in Guam in some way (perhaps by Japanese visitors choosing to reduce the length of their visit, reduce their discretionary expenditures, or consider another destination instead of Guam, which would be lower Japanese tourist arrivals in Guam, or other strategies investigated by Boone & de Hoog, 2011).

USD/JPY Trend

Figure 2 shows the trend of the USD/JPY exchange rate over the past five years, highlighting current rates of 100JPY per USD has not been experienced since early 2009.

Figure 2: Five-Year Trend in USD/JPY Exchange Rate
(Source: http://finance.yahoo.com)

Current Situation

This study is timely, given the significant strengthening of the USD/weakening of the JPY since September 2012. In the past year, Japan's central bank, i.e., the Bank of Japan (BOJ), has pursued a policy of increasing money supply in order to boost Japan's economy, which has been sluggish for 15-20 years. This policy is designed to fight the deflationary tendencies of Japan's economy by raising the inflation rate to its target of 2% per year. As a result of this policy, the U.S. dollar (USD) has strengthened and the Japanese yen (JPY) has weakened significantly from 1USD = 77.61 on September 28, 2012 to 1USD to 103.18 JPY on May 23, 2013. This represented a 33% stronger USD/weaker JPY. As noted earlier, for Japanese individuals who make purchases in USD, including those who visit Guam and other locations that use the USD as their local currency, the JPY cost had just increased 33%, even if the USD prices have not change. Since then, the Japanese yen has fluctuated around 100 JPY to 1USD,

the exchange rate that the BOJ and many Japan economy experts believe is the exchange rate that will boost domestic spending in Japan's economy sufficiently to yield a 2% inflation rate.

REVIEW OF RELATED LITERATURE

This paper investigates what factors affect tourism in general and the effect of exchange rate changes on the number of inbound tourists (or tourist arrivals in a destination country), in particular. Given the focus of this paper, the review of the literature has paid more attention on previous empirical work on the relationship between exchange rate and tourist arrival. Attempts are also made to review studies that look into tourism from and to different countries/regions in order to avoid country- or region-specific biases in tourist preferences and behaviors. Also, it should be noted that there exists no such study of this type in the context of Guam, a gap in the literature that this study is attempting to fill.

Vogt (2008) as cited in Cheng, et. al (2013, January) used annual US data from 1973 to 2002 in a partial adjustment error correction model and found that U.S. outbound tourists respond more to real exchange rate changes while U.S. inbound tourists respond more to real income changes. The opposite result (U.S. outbound tourists respond more to real income and U.S. inbound tourists respond more to real exchange rate changes) was found by Cheng, et. al (2013, January) using quarterly U.S. data from 1973 to 2010 in vector autoregressive models. Despite the opposing results found, both studies highlight the importance of two factors, real exchange rate and income, on inbound and/or outbound tourism.

Using monthly data from January 1991 and January 2011 and multivariate conditional volatility regression models, Yap (2011, March 18) investigated the effects of the appreciation of the Australian dollar on visits to Australia by tourists from nine origin countries (China, India, Japan, Malaysia, New Zealand, Singapore, South Korea, the U.K. and the U.S.) and found tourists sensitivity to stronger Australia dollar, with tourist from Malaysia and New Zealand being more sensitive. The study also found that tourists' memories of the currency changes ("shocks") could diminish in the long run, "suggesting that the sudden appreciation of Australian dollar will not have long-term negative impacts on Australia's inbound tourism".

A study by Tourism Research Australia (TRA, 2011, June) assessed the impact and relative importance of economic indicators on the travel decisions of inbound visitors to Australia. The study found that tourists' income is most important in affecting inbound tourism to Australia both in the short run and the long run, with the income elasticity of inbound tourism demand estimated as 0.8 and 1.3, respectively. As regards exchange rates, the study found that "exchange rate volatility has an impact on Australia's tourism competitiveness", with a stronger Australian dollar requiring visitors to "consider increasing their travel 'wallet' or reducing their average length of stay", with visitors still coming to Australia but making either adjustment in the short run but more likely to choose other destinations in the long run.

In response to the global economic and financial challenges since 2007, Bonner & de Hoog (2011) conducted a survey that looked at changes in the behavior of Dutch tourists, more specifically, economizing strategies they adopted in planning their vacations. Their survey included the following strategies (found on page 189 of their paper), with the top three strategies from their survey results noted:

shorter length of stay (ranked #1);
changing the destination (other country) (ranked#2);
choosing a cheaper tour operator;
choosing a self-arranged vacation instead of using a tour operator;
changing the period (earlier or later);
selecting an earlier or later booking moment;
using another means of transport;
carrying out fewer or other activities on the spot (ranked #3)
choosing another type of accommodation;
choosing a cheaper alternative within the same type of accommodation

Nowjee, et. al (2012) using a multivariate vector error correction model applied to annual data from Mauritius from 1981 to 2010 to examine the relationship between exchange rate, tourism and economic growth. Related to the present study, Nowjee, et. al (2012) found that real exchange rate did not Granger Cause tourist arrivals but found that tourist arrivals Granger cause real exchange rate.a statistically significant This means that the number of visitors to Mauritius is unaffected by changes in the exchange rate between the local currency (Mauritian rupee) and the tourists' currency. On the other hand, the number of visitors to Mauritius affects the real exchange rate, given the size of the exchange market for the Mauritian rupee and the significant size of tourism relative to the domestic economy (8.2% in 2012, Statistics Mauritius (2012)).

A study by Wang et al. (2008, November) used the Copula-based measures of dependence structure between international tourism demand and exchange rates in Asia countries constructed from available monthly data and found a negative relationship between international tourists visiting Asia and exchange rate, i.e., a stronger destination currency would reduce the number of international visitors to this destination and vice versa. The study also found an asymmetrical effect of exchange rate on international visitors, with the effect of appreciation of the destination currency stronger than the currency depreciation.

Tse (2001) estimated the impact of economic factors on tourism in Hong Kong. Measuring tourism in terms of real tourist expenditure and using an expectations model, Tse found that "real tourism expenditure depends on expected income, expected exchange rate and price level". Tse also pointed to the importance of defining the "appropriate measure of price" on international tourism. "In practice, 'price' includes the foreign currency price of tourist goods and services in destinations, the transportation cost between countries, the effect of exchange-rate variations on purchasing power. In addition, the opportunity cost of travel time and risk of travel may also be considerations." (p. 281)

Santana-Gallego, et. al (2007, December) analyzed the effect of several *de facto* exchange rate arrangements on international tourism using a gravity equation. Their findings confirm the importance of exchange rate volatility in tourists' decision to travel in that "less flexible exchange rate promotes tourism flows".

Table 4: Exchange Rate, Tourists' Incomes and Other Variables Used in Previous Studies of Tourism		
Explanatory Variables	**Author(s) & Year**	**Effect on Dependent Variable**
EXCHANGE RATE VARIABLE measured as		
Real Exchange Rate (Destination vs. Origin country Goods)	Vogt (2008)	-
Real Exchange Rate (Destination vs. Origin country Goods))	Cheng, et. al (2013, January)	-
Exchange Rate (Origin country vs. Destination country currency, Australian dollar)	Yap (2011)	- but diminishes in the long run
Exchange rate elasticity of international tourism demand	TRA (2011, June)	- with differential adjustments in the short run vs. long run
Real exchange rate (Origin country vs. Destination country, Mauritian, Goods)	Nowjee, et. al (2012, November)	no effect
Exchange rate (Foreign currency vs. Destination country (select Asian country) currencies)	Wang, et. al (2008, November)	- with asymmetrical response, i.e., stronger sensitivity to domestic currency appreciation than depreciation
Expected Exchange Rate	Tse (2011)	
Exchange rate volatility (proxy for *de facto* exchange rate arrangements)	Santana-Gallego, et. al (2007)	-
Exchange rate elasticity of inbound tourism	Crouch (1993)	-
OTHER VARIABLES measured as		
Income	Vogt (2008) used Real Income	+
	Cheng, et. al (2013, January) used Real Income	+
	TRA (2011, June) used Income elasticity of international tourism demand	+ with short-run being more elastic than long-run
	Tse (2011) used Expected Income	
	Crouch (1993) used Income	+
Price Level	Tse (2011)	
Time Period being analyzed	Crouch (1993)	
Relative inflation rates (Origin vs. Destination country inflation rates)	Crouch (1993)	+

The paper by Crouch (1993, December) reviewed empirical studies to-date on the impact of exchange rates on international tourism demand and found the impact to be significant but noted the large variability on the estimates of this impact found by these studies. Using a meta-analytical approach, he then investigated this variability among 286 exchange rate elasticities of demand from 80 empirical studies and found the importance of including (1) tourists' income in the model along with exchange rate because "as the currency of the origin country drops in value, the standard of living and real incomes normally decline. The decline in income and the increase in exchange rates together deter foreign tourism"; (2) relative inflation rates, arguing that "as the currency of the origin country drops in value, inflation normally increases. The price of a destination in the form of relative rates of inflation might therefore decline," and, related to the finding of Santana-Gallego, et. al (2007, December), Crouch found that a change in exchange rate systems might affect trend in exchange rate elasticities of international demand.

Table 4 summarizes the results of those studies just reviewed with regard to the effects of exchange rate and other variables on tourism while Table 5 identifies the dependent variables, time periods and origin/destination countries used in the studies just reviewed.

Table 5: Dependent Variable, Time Period, and Country Groups in Previous Studies

Author(s) & Year	Dependent Variable, Time Period, Origin/Destination Countries
Vogt (2008)	Exports revenue to U.S., 1973-2002 quarterly data
Cheng, et. al (2013, January)	Tourist arrivals to/from eight Asian countries (Japan, China, Korea, Taiwan, Hong Kong, Singapore, Malaysia and Thailand, January 2001-July 2007 monthly data
Yap (2011)	Tourist arrivals to Australia from China, India, Japan, Malaysia, New Zealand, Singapore, South Korea, the UK and the USA, January 1991-January 2011 monthly data
TRA (2011, June)	Tourist arrivals to Australia, 1990-2010 data frequency unknown
Nowjee, et. al (2012, November)	Tourist arrivals to Mauritius, 1981-2010 annual data
Wang, et. al (2008, November)	Exports revenue to U.S., 1973-2010 quarterly data
Tse (2011)	Tourist arrivals to and hotel room rates in Hong Kong, 1973-1998 annual data
Santana-Gallego, et. al (2007)	Log of tourist arrivals to multiple countries grouped according to *de facto* exchange rate regimes , 1995-2001
Crouch (1993)-survey of previous studies	Tourist arrivals, tourist expenditures, multiple time periods and origin and destination countries

EMPIRICAL MODEL OF JAPANESE TOURISM IN GUAM

Given Guam's heavy reliance on Japanese visitors and the significantly stronger USD/weaker JPY in the past year, which for Japanese visitors makes a visit to Guam more expensive, this study uses Ordinary Least Squares regression analysis and monthly data from October 2003 to July 2013 (a period of 115 months) to measure the effect of a stronger USD/weaker JPY on the number of Japanese tourists visiting Guam.

The Empirical Model

In this study, the regression equation is

$$\begin{aligned}&\textbf{Japanese Tourist Arrival in Guam}_t = \\ &a_0 + a_1\,\textbf{USD/JPY}_{t-i} + a_2\,\textbf{Japanese Growth}_t + a_3\,\textbf{Tohoku Disaster} + \\ &a_4\,\textbf{Trend} + a_5\,\textbf{Monthly Seasonality} + a_6\,\textbf{Japanese Tourist Arrival in Guam}_{t-1} + e_t\end{aligned} \quad (1)$$

where the dependent variable is **Japanese Tourist Arrival in Guam$_t$** = number of Japanese tourists arriving in Guam in month t. This variable is consistent with the dependent variables used by several studies in Table 5. Data was taken from various issues of Guam Visitors Bureau's Visitor Arrivals Statistics.

The independent/explanatory variables in the regression equation are

USD/JPY$_{t-i}$ = Nominal exchange rate between JPY and USD (how many JPY is required to buy 1USD) at time t-i, where i= 1 to 12 to indicate 1 to 12 month lagged effect of exchange rate. Data was downloaded from the Federal Research Bank of St. Louis, Federal Reserve Economic Data (FRED2), series ID: EXJPUS.

Japanese Growth$_t$ = Growth of Japanese tourists' real income, proxied by Japan's monthly industrial production index, which was downloaded from the Federal Research Bank of St. Louis, Federal Reserve Economic Data (FRED2), series ID: JPNPROINDMISMEI (2005=100).

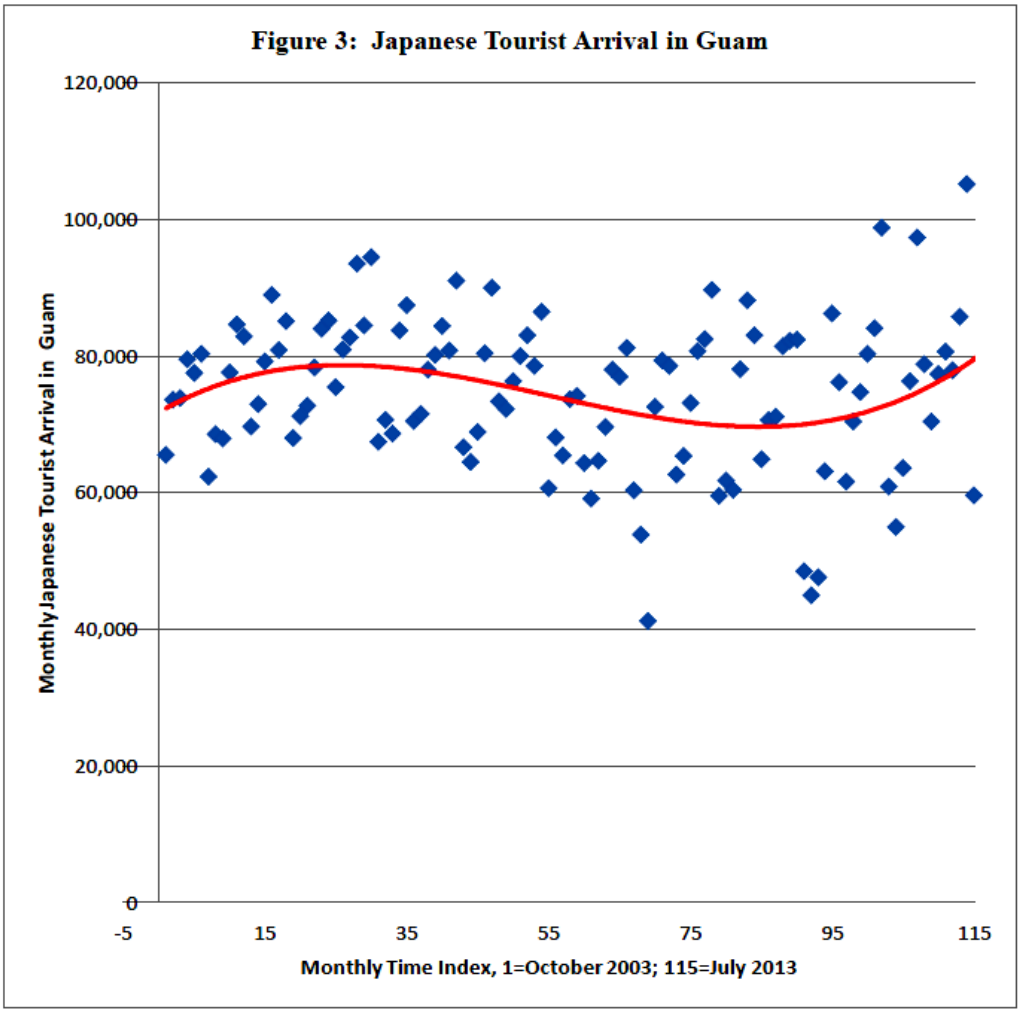
Tohoku Disaster = dummy for the March 2011 earthquake and tsunami disaster in northeastern Japan (Tohoku area), which noticeably reduced the number of Japanese visitors to Guam in the three months following the disaster, i.e., April, May and June 2011.

Trend = index for months of time series data, from 1=October 2003 to 115=July 2013. Figure 3 shows the trend of the dependent variable (Japanese Tourist Arrival in Guam) to mimic a cubic function.

Monthly Seasonality = dummy for the monthly seasonality in the dependent variable (Japanese Tourist Arrival in Guam). Figure 3 shows monthly seasonality around the cubic trend displayed by Japanese Tourist Arrival in Guam. A separate regression analysis shows particular seasonality for the months of January, February, March, April, May, June, August and October compared to the month of December. On the other hand, the months of July, September and November did not show significantly different seasonality than that for the month of December.

Japanese Tourist Arrival in Guam$_{t-1}$ = introduced to capture any autoregressive pattern of the dependent variable

The error term is indicated by **e_t** in the regression equation.



Source: Guam Visitors Bureau (various issues). Visitor Arrivals Statistics. Retrieved from http://www.visitguam.org

The regression equation in (1) is estimated using Ordinary Least Squares and processed using Microsoft Excel/Data Analysis/Regression.

The Test Hypotheses

The empirical model will test the following hypotheses:

H1: A stronger USD/weaker JPY will negatively affect Japanese tourist arrival in Guam ($a_1<0$).

H2: Higher Japanese tourists' income will positively affect Japanese tourist arrival in Guam ($a_2>0$).

H3: The Tohoku disaster in March 2011 has negatively affected Japanese tourist arrival in Guam ($a_3<0$).

H4: Japanese tourist arrival time series displays a cubic trend with respect to time (in this case, months) (a_4 >0 for Trend, <0 for Trend^2 and >0 for Trend^3).

H5: Japanese tourist arrival time series displays monthly seasonality with some months experiencing stronger Japanese tourist arrival and other months experiencing weaker Jpaanese tourist arrival than the reference month (December) (a_5 >0 for months stronger than the reference months; a_5 <0 for months weaker than the reference months).

H6: Lagged Japanese tourist arrival in Guam positively affects current Japanese tourist arrival in Guam (a_6>0).

THE RESULTS

Multiple regression runs were performed in order to identify the effect of the USD/JPY exchange rate of different lags (from one month to twelve months) on the Japanese Tourist Arrival in Guam, as reflected by coefficient $\mathbf{a_1}$ in the regression equation in (1). This study hypothesized $\mathbf{a_1}$ to be negative. Estimates of $\mathbf{a_1}$ for different lags on the USD/JPY are reported in Table 6.

Table 6: Effect of the Stronger USD/Weaker JPY on Japanese Tourist Arrival in Guam

Time Lag (in months)	Estimated value of $\mathbf{a_1}$	p-value	significance
0 (current time)	-159.72	0.1116	None
1	-176.42	0.0786	*
2	-199.98	0.0383	**
3	-152.89	0.1009	None
4	-76.52	0.4003	None
5	-47.51	0.5976	None
6	-84.17	0.3404	None
7	-120.88	0.1644	None
8	-110.9	0.1967	None
9	-163.99	0.0527	*
10	-178.42	0.0326	**
11	-146.38	0.0761	*
12 (1 year earlier)	-136.51	0.0955	*

* indicates a 10% significance level
** indicates a 5% significance level
otherwise, the coefficient is not significantly different from zero

The results reported in Table 6 show negative values of $\mathbf{a_1}$ for USD/JPY for the following lags in months: 2, 3, 9, 10, 11 and 12. The magnitude of this effect ranges from -135.51 using a 12-month lag on the USD/JPY exchange rate to -199.98 using a 3-month lag. These estimates

are to be interpreted as representing the reduction in the number of Japanese tourists visiting Guam per month for every 1JPY that the JPY is weaker vs. the USD. As such, these effects are significant since the JPY has weakened from 76JPY to 100JPY per 1USD, or 24JPY per 1USD, since September 2012. This means that during this period, the estimated reduction in Japanese tourists visiting Guam ranges from 3,252 to 4,799 per month or 39,026 to 57,594 over a 12-month period, which translates to a decline of between 4.25% and 6.27% in the number of Japanese tourist expected to visit Guam during this current fiscal year (October 2012-September 2013).

The economic impact of this estimated reduction in Japanese visitor to Guam in response to the unfavorable exchange rate faced by Japanese tourists would be quite noticeable, especially if not offset by positive contributions by visitors to Guam from other countries. These estimates are calculated using the same methodology presented earlier, which focused on tourist expenditure in Guam. With each Japanese visitor spending in Guam almost USD500 during his/her visit to Guam and expecting between 39,026 to 57,594 less Japanese tourists to visit Guam in fiscal year 2013, this would

reduce tourist expenditure by between USD19.5 million to USD28.8 million
reduce the overall Guam economy by the spending multiplier of 1.3 (approximately between USD25.4 million to USD37.4 million)
reduce the number of jobs by between 346 and 512; and
reduce taxes in the form of the Gross Receipts Tax (GRT) by between USD1 million to USD1.5 million,
and other negative economic impacts not included here because of their calculations would require information beyond what is obtained for this study.

The results reported in Table 6 also suggest that the negative effect of the stronger USD/weaker JPY on the number of Japanese tourists arriving Guam appears to be experienced in the short-run (in this case, 2-3 months after the change in the USD/JPY exchange rate) and later, in the long-run (from 9 to 12 months after the change in the USD/JPY exchange rate). The latter is consistent with those Japanese tourists who make early travel plans (up to one year in advance; Schumann, F.R., 2013, May, personal communication), many of whom book packaged tours (25% booked "full tour packages" while 68% booked "free-time package tours", Guam Visitors Bureau's Japan Visitor Tracker Exit Profile, June 2013, prepared by Qmark Research). The former likely reflects those Japanese tourists who make late travel plans, which they booked themselves (referred to as "individually arranged travel", which accounted for 4% of the respondents to Guam Visitors Bureau's Japan Visitor Tracker Exit Profile, June 2013, prepared by Qmark Research.

Other explanatory variables were also found to have statistically significant effects on the dependent variable, Japanese Tourist Arrival in Guam. As mentioned earlier, multiple regression runs were processed. Tables 7 and 8 report the regression results where the coefficient $\mathbf{a_1}$ has the lowest p-values, which according to Table 6 were those with the USD/JPY exchange rate with a 3-month lag as well as a 10-month lag.

Results reported in Tables 7 and 8. All coefficients were found to be statistically significant at a 1% level, except for **USD/JPY$_{t-i}$** (where i=2 and 10) and **Japanese Tourist Arrival in Guam$_{t-1}$**, which were statistically significant at a 5% level. The R^2 and adjusted R^2 are high (low to mid-80%) and the F-statistics are statistically significant at a 1% level or better, as shown by extremely low p-values.

We reiterate that a stronger USD/weaker JPY reduces the number of Japanese tourists arriving in Guam, a result that was already discussed and for which estimated coefficients corresponding to different time lags were presented in Table 6

Table 7: OLS-Regression Results, USD/JPY exchange rate lagged 2 months

Dependent Variable=Japanese Tourist Arrival in Guam$_t$ (n=115)

Explanatory Variables ↓	**Coefficients**	**Standard Error**	**t Stat**	**P-value**
Intercept	81020.56	10596.64	7.65	1.37E-11
USD/JPY$_{t-2}$	-199.98	95.23	-2.10	0.038384
Japanese Growth$_t$	18170.64	4683.22	3.88	0.000188
Dummy for Tohoku Disaster	-9292.52	3095.51	-3.00	0.003396
Trend	826.78	218.90	3.78	0.000271
Trend^2	-19.16	4.59	-4.17	6.48E-05
Trend^3	0.11	0.02	4.42	2.57E-05
Monthly Seasonality Dummy: January	6152.82	1740.65	3.53	0.000622
Monthly Seasonality Dummy: February	3596.54	1852.25	1.94	0.055015
Monthly Seasonality Dummy: March	11773.51	1818.03	6.48	3.66E-09
Monthly Seasonality Dummy: April	-16451.20	2105.13	-7.81	5.97E-12
Monthly Seasonality Dummy: May	-12492.70	1883.26	-6.63	1.75E-09
Monthly Seasonality Dummy: June	-11298.30	1916.84	-5.89	5.22E-08
Monthly Seasonality Dummy: August	9519.77	1807.45	5.27	8.1E-07
Monthly Seasonality Dummy: October	-8926.82	1818.91	-4.91	3.63E-06
Japanese Tourist Arrival in Guam$_{t-1}$	0.1480	0.067	2.22	0.028755
	R^2	**0.8419**	**F-statistics**	**35.1434**
	Adjusted R^2	**0.8179**	**P-value of F**	**6.61E-33**

We also find that an increase in Japanese tourists' real income, as proxied by growth in Japan's monthly industrial production, encourages visits to Guam, as reflected by a positive estimated for $\mathbf{a_2}$ of approximately 18,000. This means that, for every one-percentage point increase in real income of Japanese tourists, an additional 18,000 Japanese tourists will visit Guam. This result suggests that, Japanese tourists view visiting Guam as a normal good.

Our empirical model also captures the negative impact of the earthquake and tsunami disaster that affected northeastern Japan on March 11, 2011 on the number of Japanese visitors arriving in Guam. The estimates for $\mathbf{a_3}$ of between 9,300 (Table 7) and 10,000 (Table 8) correspond to the reduction in the number of Japanese visitors to Guam during the months of April, May and June, 2011. Figure 3 also clearly shows the data points corresponding to these months to be outliers and significantly below the cubic trend line.

As shown in Figure 3, our regression results confirm that the Japanese tourist arrival time series data exhibits a cubic function with respect to its monthly trend, as reflected in the estimated coefficients for Trend, $Trend^2$ and $Trend^3$.

As also evident in Figure 3, we find that there are monthly seasonality in Japanese tourist arrival in Guam, with March being the busiest month and representing the highest arrivals, followed by August, then January and February and all these months outperforming the months of July, September, November and December. April was found to be the slowest month in terms of Japanese tourist arrival in Guam, followed by May, June and October, with these months corresponding to Japanese tourist arrival in Guam to be lower than those during the months of July, September, November and December.

Table 8: OLS-Regression Results, USD/JPY exchange rate lagged 10 months

Dependent Variable=Japanese Tourist Arrival in $Guam_t$ (n=115)

Explanatory Variables ↓	**Coefficients**	**Standard Error**	**t Stat**	**P-value**
Intercept	81888.82	10688.12	7.66	1.26E-11
USD/JPY_{t-10}	-178.46	82.31	-2.17	0.032583
Japanese $Growth_t$	16046.29	4603.28	3.49	0.000733
Dummy for Tohoku Disaster	-9976.42	3114.21	-3.20	0.001827
Trend	594.39	161.13	3.69	0.000368
$Trend^2$	-14.01	3.24	-4.33	3.6E-05
$Trend^3$	0.0779	0.02	4.40	2.72E-05
Monthly Seasonality Dummy: January	6581.05	1739.82	3.78	0.000266
Monthly Seasonality Dummy: February	4366.27	1868.29	2.34	0.021452
Monthly Seasonality Dummy: March	12485.01	1826.35	6.84	6.75E-10
Monthly Seasonality Dummy: April	-15659.10	2140.74	-7.31	6.82E-11
Monthly Seasonality Dummy: May	-12091.00	1887.05	-6.41	5.03E-09
Monthly Seasonality Dummy: June	-11042.40	1916.31	-5.76	9.4E-08
Monthly Seasonality Dummy: August	9464.63	1804.41	5.24	8.88E-07
Monthly Seasonality Dummy: October	-8987.42	1816.84	-4.95	3.09E-06
Japanese Tourist Arrival in $Guam_{t-1}$	0.1369	0.067	2.04	0.043966
	R^2	**0.8423**	**F-statistics**	**35.2602**
	Adjusted R^2	**0.8184**	**P-value of F**	**5.78E-33**

Our empirical model finds that, on average, Japanese tourist arrival in Guam in any particular month is positively affected by arrival during the previous month, as indicated by the estimated for $\mathbf{a_6}$ of 0.148 (Table 7) and 0.1369 (Table 8).

CONCLUSION AND POLICY IMPLICATIONS

This study aimed at investigating the relationship between exchange rates and tourism using evidence from the Guam economy. Our empirical model confirms that a stronger USD/weaker JPY would discourage Japanese visitors to Guam. The combination of Guam's heavy reliance on the Japanese tourist market, which accounts for approximately 70% of tourist arrival in Guam, the relatively large amount of expenditure by Japanese tourists while in Guam (approximately USD500 per tourist per visit) and the 33% strengthening of the USD vs. the JPY in the past year point to the noticeably large impact on Guam's USD 4 billion economy in terms of the reduced overall income and spending, employment and tax collections by the local governments. Although the exchange rate appears to have stabilized around JPY100 per 1USD, which represents a preliminary target by the Bank of Japan, the worst might not be over since this preliminary target was believed to bring Japan's inflation rate to 2%. As Japan's inflation rate continues to fall below 2%, which reflects continued slow economy and tendencies of deflationary pressures, the possibility remains for another round of JPY depreciation in order to encourage exports from Japan in the hopes that this would boost the sluggish Japanese economy. In this scenario, further weakening of the JPY would mean further strengthening of the USD, which would increase the costs to Japanese tourists of visiting Guam.

On the other hand, to the extent that the further weakening of the JPY would stimulate the Japanese economy, incomes of Japanese tourists would increase, which would create additional purchasing power for Japanese consumers and encourage visits to Guam. Our findings suggest that the exchange rate effect would become visible first, as early as two months after another exchange rate adjustment and certainly within the first year of the adjustment.

Despite what continues to be a heavy reliance of Guam's tourism on the Japanese market, the fact is that the share of Japanese visitors to the total has been reduced to approximately 70% from what was much higher (85-90%), thanks to many years, even decades, of efforts by the Guam Visitors Bureau and its members to diversify Guam's tourism by proactively marketing to other tourist markets. Also contributing to this change are market and institutional factors that increase Guam's accessibility and affordability to tourist from other origin countries. Visitors to Guam from Korea now make up 17.45% of the total, with the Guam Visitors Bureau's plan to increase this figure to 30% in the near future. Only 5 years ago, the share of Korean tourist was as low as 12-13%. Of course, the increased share of Korean tourist also resulted from the weakened U.S. dollar vis-à-vis the Korean won during the same period (Cruz, B.J., 2013, October 3, personal communication). Russia's small share (0.51%) to the total tourist arrival in Guam brings promise of triple-digit growth for some time to come. Fortunately, this growth

prospect is driven largely by pent-up demand for travel by Russian tourists and likely to be immuned from the state of the Russia economy or the exchange rate between the USD and the Russian Ruble for some time to come. Another market with a lot of promise for Guam's tourism is Mainland China, which make up only 0.79% of Guam's tourist market. To this end, there continues to be efforts by the Guam Visitors Bureau and some local policymakers to push to include China in the Visa Waiver Program. These and other efforts combine to offset the negative impact of the stronger USD/weaker JPY on Japanese tourist arrival in Guam. Based on the latest figures, that the overall tourist arrival in Guam manages to increase 7% this fiscal year suggests that these efforts have been effective.

REFERENCES

Bonner, Fred & Robert de Hoog (2011). Economizing behavior during travel: Strategies and information sources used. *Journal of Vacation Marketing 17(3)*, 185–195.

Cheng, Ka Ming, Hyeongwoo Kim, & Henry Thompson (2013, January). The real exchange rate and the balance of trade in US tourism. *International Review of Economics and Finance 25*, 122-128.

Crouch, Geoffrey I. (1993, December). Currency exchange rates and the demand for international tourism. *The Journal of Tourism Studies 4(2)*, 45-53.

Federal Reserve Bank of St. Louis, Federal Reserve Economic Data, http://research.stlouisfed.org/

Guam Department of Labor-Bureau of Labor Statistics (2013, June). *Current Employment Statistics.*

Guam Visitors Bureau (various issues). Visitor Arrivals Statistics. Retrieved from http://www.visitguam.org

Guam Visitors Bureau (2013, April-June). Hong KongVisitor Tracker Exit Profile, prepared by Qmark Research. Retrieved from http://www.visitguam.org

Guam Visitors Bureau (2013, June). Japan Visitor Tracker Exit Profile, prepared by Qmark Research. Retrieved from http://www.visitguam.org

Guam Visitors Bureau (2013, July). Korea Visitor Tracker Exit Profile, prepared by Qmark Research. Retrieved from http://www.visitguam.org

Guam Visitors Bureau (2013, January-March). Russia Visitor Tracker Exit Profile, prepared by Qmark Research. Retrieved from http://www.visitguam.org

Guam Visitors Bureau (2013, April-June). Taiwan Visitor Tracker Exit Profile, prepared by Qmark Research. Retrieved from http://www.visitguam.org

Nowjee, A., V. Poloodoo, M. Lamport, K. Padachi, & D. Ramdhony (2012). The relationship between exchange rate, tourism and economic growth: Evidence from Mauritius. Conference Proceedings from the International Conference on International Trade and Investment (ICITI) 2012.

Ruane, Maria Claret M. (2011, December). Estimating the spending multiplier on Guam. University of Guam-Pacific Center for Economic Initiatives Technical Report, Issue 15.

Santana-Gallego, Maria, Francisco J. Ledesma-Rodriguez, & Jorge V. Perez-Rodriguez (2007, December. On the impact of exchange rate regimes on tourism. *Documentos de Economia y Finanzas Internacionales DEFI 07-07*, 1-16.

Statistics Mauritius, Government of Mauritius-Ministry of Finance and Economic Development Department (2012). *National Accounts of Mauritius 2012*. Retrieved on September 10, 2013, from http://statsmauritius.gov.mu/English/Publications/Documents/Regular%20Reports/national%20accounts/narreport2012.pdf

Tourism Research Australia, Australia Government-Department of Resources, Energy and Tourism (2011, June). *Factors affecting the inbound tourism sector – the impact and implications of the Australian dollar*.

Tse, Raymond Y.C. (2001). Estimating the impact of economic factors on tourism: Evidence from Hong Kong. *Tourism Economics 7(3)*, 277-293.

U.S. Census Bureau (2010). Guam Population Estimates.

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.

Vogt, M.G. (2008). Determinants of the demand for US exports and imports of tourism. *Applied Economics 40*, 667-672.

Wang, Hui-Cheng, Nai-Hua Chen, Ching-Lung Lu, Tsorng-Chyi Hwang, & Shuo-Wen Tseng (2008, November). Tourism demand and exchange rate in Asian countires: New evidence from Copulas approach. *Convergence and Hybrid Information Technology 2*, 1188-1193.. ICCIT '08. Third International Conference.

Yahoo! Finance, http://finance.yahoo.com

Yap, Ghialy C. L. (2011, March 18). Modelling the spillover effects of exchange rates on Australia's inbound tourism growth. Available at SSRN: http://ssrn.com/abstract=1789645 or http://dx.doi.org/10.2139/ssrn.1789645

RECENT TRENDS AND NEW EVIDENCE IN ECONOMICS LITERACY AMONG ADULTS

Celeste Varum, University of Aveiro, GOVCOPP
Eduarda Santos, University of Aveiro
Vera Afreixo, University of Aveiro, CIDMA

ABSTRACT

Economics literacy has received growing attention in the academic literature and even more in the context of the present economic and financial crisis. In this work we develop a sound and novel empirical work, analysing the level and determinants of economic literacy of a sample of adults in Portugal, being unique in this respect. The purpose of this paper is to obtain new evidence about a fundamental question of empirical studies on economic literacy: the determinants of the level of economic literacy. Besides, we investigate the level of economic literacy of adults and interest on the matter. A good evaluation of economic literacy allows one to distinguish the existing deficiencies and thus define education according to these deficiencies. It is expected that this work will contribute to an increased interest in "education in economics" on the part of researchers and that their results will allow for the expansion of knowledge about the Portuguese reality, being possible to compare the results to others obtained internationally. The developed questionnaire can also be applied by other researchers in the future.

INTRODUCTION

It is essential for citizens to have a reasonable level of knowledge about the functioning of the economy lato sensu, or relative to markets of goods and services, work, and capital, in a society that intends more active citizen participation. An understanding of market functioning will make it possible for citizens to evaluate political decisions and their consequences in a more fundamental manner, as well as make better decisions that maximise their well being (Koshal, Gupta, Goyal, & Choudhary, 2008). Huston (2010) and Remund (2010) concludes that it is extremely important to increase the general level of the population's economic knowledge so that people can better understand and settle the decisions with which they are currently confronted. Economics literacy has received growing attention in the academic literature and even more in the context of the present economic and financial crisis.

Not surprisingly, economics literacy, which encompasses both real and financial aspects, has received growing attention in the academic literature (Clark, Shung & Harrison, 2009), and

even more in the context of the present economic and financial crisis. In this context, it becomes relevant to further investigate the citizens' level of economic knowledge, as well as to explore variables that permit explanations of differentiation between individuals' economic knowledge.

In this work we develop a sound and novel empirical work, analysing the level and determinants of economic literacy of a sample of adults in Portugal, being unique in this respect. A good evaluation of economic literacy allows one to distinguish the existing deficiencies and thus define education according to these deficiencies. It also permits identification of the more critical groups (Huston, 2010). The vast majority of this theoretical and empirical literature focuses on the USA case and emphasises financial aspects, but interest in this subject continues to gain interest and attention from researchers, teachers, institutions and political decision-makers in other parts of the world. This study analysis new data for an European economy, Portugal. It is expected that this work will contribute to an increased interest in "education in economics" on the part of researchers and that their results will allow for the expansion of knowledge about the Portuguese reality, being possible to compare the results to others obtained internationally. The developed questionnaire can also be applied by other researchers in the future.

BACKGROUND

Economic literacy consists of the set of knowledge and competencies that permit the improvement of personal and social decisions about various economic problems encountered in daily life, whether as consumers, vendors, producers, investors, workers or voters. An important component of economic literacy involves knowledge of financial aspects or financial literacy. Financial literacy is understood as the comprehension of a set of economic concepts that can be used to evaluate financial situations and make good financial decisions (Pang, 2010).

It becomes relevant to investigate the citizens' level of economic knowledge, as well as to explore variables that permit explanations of differentiation between individuals' economic knowledge. On this regard, studies in the literature reveal the importance of the education level. Gleason & Scyoc (1995), Wood & Doyle (2002) and Walstad & Rebeck (2002) verified that the education level of individuals had a statistically significant positive effect on their economic literacy, such that the greater the level of education, the greater the level of correct responses on a test on economics. More recently, Monticone (2010) verified that the highest education levels are generally associated with higher levels of financial knowledge. Individuals with more education experience fewer difficulties when acquiring financial knowledge and therefore incur fewer learning costs. In an international comparison, Jappelli (2010) verified that, at country level, the general level of education is positively related to the level of economic literacy.

The literature also indicates that having training or a degree in economic sciences is also important to possessing economic knowledge. Wood & Doyle (2002) and Koshal, Gupta, Goyal

& Choudhary (2008) verified that possession of a degree in economic sciences has a positive effect on economic literacy. Soper & Brenneke (1981), Gleason & Scyoc (1995) and Walstad & Rebeck (2002), and Walstad and Rebeck (1999), also concluded that adults with degrees and/or training in economics know more about economics than those that do not have training in economics. From a complementary perspective, Walstad, Rebeck & MacDonald (2010) investigated whether training in personal finances during secondary education increased the level of financial knowledge. The authors verified that the levels of financial knowledge increased significantly in the students who participated in a personal finance training. In this line, Pang (2010) published a study in which a specialised course was applied to increase financial literacy to students in secondary education and thus enable them to make informed and independent financial decisions. The results showed that the students who attended the course performed better than those who did not attend the course and that this advantage was maintained over time. Income level is another factor that is highlighted in the literature. In a study by Monticone (2010), the connection between financial behaviour and financial knowledge was studied with a focus on the accumulation of wealth. The results indicated that families with greater wealth had a greater probability of investing in financial knowledge. Wood & Doyle (2002), Walstad & Rebeck (2002) and Grimes, Millea & Thomas (2010) also verified that economic knowledge is consistently affected by the income level, thus confirming that individuals with higher salaries possess more economic knowledge. In a study by Jappelli (2010), in which an international comparison is made between different countries, it was verified that economic literacy tends to be associated with higher incomes. However, Mandell & Klein (2007) concluded that family income is not a determinant of financial literacy.

Some studies explore the effect of gender on economic literacy. A significant part of this literature indicates that, on average, males have consistently higher levels of economic knowledge than females (for example, Soper & Brenneke, 1981; Gleason & Scyoc, 1995; Walstad & Rebeck, 2002; Wood & Doyle, 2002; Tabesh & Schultz, 2007;, Millea & Thomas, 2010; Monticone, 2010).

Nonetheless, a few have concluded that gender does not influence literacy levels (see Mandell & Klein, 2007; Koshal, Gupta, Goyal & Choudhary, 2008).

In the literature, it is reported that individuals learn economics throughout their lifetimes (Grimes, Millea & Thomas, 2010). Thus, naturally, age has been indeed considered to be a determining factor in economic literacy in several studies, such as Gleason & Scyoc (1995) and Walstad & Rebeck (2002). However, age may not have a linear relationship with learning. For this reason the authors often test the effect of age squared (Walstad & Rebeck, 2002; Koshal, Gupta, Goyal & Choudhary, 2008; Monticone, 2010). Indeed, Koshal, Gupta, Goyal & Choudhary (2008) verified that the economic literacy of MBA students increased with age, although at a decreasing rate. Walstad & Rebeck (2002) and Monticone (2010) reported that the

relation between age and financial knowledge is an inverted U (concave), which means that middle-aged adults have higher levels of literacy than those who are younger and older. Both studies verified that literacy increases until 40-60 years of age and then declines because knowledge only accumulates until a certain age and later depreciates. This phenomenon might also occur because older generations were not exposed to the current complex financial services during their youth.

The effect of an individual's professional situation on financial literacy as also been analysed. Monticone (2010) confirmed that employed individuals responded correctly to more questions than did those who were unemployed or out of the work force.

Another factor analysed in the literature is ethnicity. Both studies by Mandell & Klein (2007) and Grimes, Millea & Thomas (2010) concluded that ethnicity was a determinant in literacy by verifying that Africans had lower literacy levels. Experience in the job market was analysed in a study by Koshal, Gupta, Goyal & Choudhary (2008). Because experience generally does not have a linear relationship with learning, the authors included the squared number of years of experience in the job market. They verified that the marginal rate of economic literacy on the order of experience increased at a greater rate, which suggested that gains in economic literacy are accelerated by experience in the job market.

Finally, one more factor was analysed in the literature, mathematics knowledge. Mathematics knowledge was confirmed to have a statistically significant positive effect on economic literacy (Jappelli, 2010; Schuhmann, McGoldrick, & Burrus (2005).

In the next section we address the following issues. What is the general knowledge of economics in the adult community? Are adults capable of understanding economic and financial concepts? Which factors explain the differences in levels of economic literacy among the community in general?

METHODOLOGY AND DATA

The questionnaire designed draws predominantly on the Economy Literacy Test (ELT) developed by the NCEE because there was no standardised tool with which to evaluate the economic literacy of Portuguese adults. The ELT was chosen because its reliability, validity and consistency have been proven over a 13-year period by thousands of respondents.

The possibility of evaluating the financial knowledge of adults in the same questionnaire also emerged since, given the international financial crisis, financial concepts are extremely important to the population. A questionnaire that was applied by the Bank of Portugal to the Portuguese population in 2010, known as the "Inquiry of Financial Literacy of the Portuguese Population" ["Inquérito à Literacia Financeira da População Portuguesa" (ILFPP)], was used. This questionnaire was chosen because it was already adapted to the economic and financial realities of the Portuguese population and thus did not require either translation or adaptation of

international terms. Thus, the questionnaire used in this study to measure the economic literacy and financial comprehension of adults in the general population combined questions from the Economic Literacy Test and from the ILFPP of the Bank of Portugal. In total, the study questionnaire has 29 questions, of which 22 address economic questions and 7 address financial questions that allowed us to assess the financial comprehension of adults as shown in Table 8. Sociodemographic, economic and motivational variables relative to the degree of interest, attitudes, ambition and importance of economics to each individual were also collected.

The questionnaire is subdivided into four parts. The first part comprises a set of questions regarding the sociodemographic characteristics of the respondents. Part two includes multiple-choice questions that evaluate the economic literacy of the respondents. The questionnaire used in this study includes questions that were included in the 1999 and 2005 versions of the ELT as well as financial questions that were included in the ILFPP, which was applied to the Portuguese population by the Bank of Portugal in 2010. It was considered relevant to add two more questions about International Economics because the original questionnaire (ELT) only included two questions related to this subject. It comprises 22 questions that address the following economic areas: consumer economics, producer economics, financial economics, the economic role of the government and international economics.

The part 3 includes seven multiple-choice questions that evaluate the respondent's financial comprehension of basic financial concepts, which are encountered by a significant portion of the population in daily life. In this group, the respondent is also asked about the income class in which their household monthly income is located.

Finally, part 4 attempts to analyse the respondents' interest in economics and the degree of importance that it has on economic subjects. Thus, the respondents are questioned about whether they follow economic subjects or news through various means of communication and whether economic knowledge is important to the perception of electoral promises, to being a more responsible citizen, to making better investment decisions and to improving well-being. Respondents are also asked if they are able to save and what their main motives for saving are. Finally, the respondents reveal their degree of interest in economic subjects and whether they consider it relevant to insert economic subjects into basic education programs for students. This last group will be relevant when explaining that, in addition to the economic and demographic characteristics influencing the levels of economic and financial literacy, motivation and an interest in economics also significantly influence the levels of economic and financial literacy.

The degree of internal consistency in the questionnaire was assessed with the Cronbach alpha coefficient. This coefficient varies from 0 to 1, and the greater the value of this coefficient, the greater the consistency and reliability of the questionnaire. According to various authors (e.g., Belbute & Sousa (2004)) in inquiries with elevated numbers of questions, an alpha value greater than 0.7 shows a good level of internal consistency and reliability (Nunnaly, 1978). In

this study, the value achieved by this coefficient was 0.902, which permits the conclusion that the questionnaire is reliable; this was expected because the questionnaire was developed from an existing and tested questionnaire.

The data used were collected with a questionnaire that was applied in April 2012 to the parents/guardians and teachers of students who were attending the 1st Cycle of Basic Education in the Aveiro Schools (5 schools). The questionnaire performance implies the choice of a sample that reflects, in an unbiased manner, the characteristics of the universal population such that it is possible to use answers from the respondents to estimate, through statistical inference, the degree of economic and financial literacy of adults in general. A total of 1061 questionnaires (1016 parents and 45 teachers) were distributed. Out of the 1061 delivered questionnaires, only 618 properly filled-out questionnaires were returned, of which 598 were from parents/guardians and 20 were from teachers. Thirty-seven blank questionnaires and 2 incomplete questionnaires were excluded. The largest percentage of the collected questionnaires are from school C (35.8%) followed by A (29.9%) and B (22.2%). Overall, 96.6% of the respondents were parents/guardians. The responses from the teachers corresponded to only 3.4% of the total (Table 12).

Table 1. Stratification of the collected sample relative to the demographics of the respondents (%)

Individual characteristics		**(%)**
Gender	Female	70,9
	Male	28,4
	N.reply	0,6
Nacionality	Portuguese	93,2
	Other	6,6
	N.reply	0,2
Age	26-35	14,9
	36-45	67,7
	46-55	15,5
	56-67	1,5
	Não responde	0,5
Civil status	Married	68,2
	Union	9,5
	Single	6,5
	Divorced	13,4
	Widow	1,1
	Not reply	1,3

As shown in Table 1, the sample was found to be mostly composed of females (70.9%). Additionally, the majority of the respondents were found to be Portuguese (93.2%). It was

observed that the respondents' ages varied between 26 and 67 years. The majority of the respondents were married (68.2%).

Table 2. Stratification of the collected sample relative to the qualifications and capacities of the respondents (%)

		%
Education level	No primary school	0,3
	4 years	4,0
	9 years	11,5
	12 years	24,4
	University degree	43,3
	Post graduation	15,3
	No reply	1,1
Holds degree in Economics/ finance	Yes	13,9
	No	64,5
	No reply	21,6
Holds training in economic related subjects	Yes	26,3
	No	62,4
	No reply	11,3
Professional activity (1: professions associated to higher qualifications; 9-professions associated to less qualifications)	Group 1	4,4
	Group 2	37,8
	Group 3	11,8
	Group 4	12,1
	Group 5	10,2
	Group 6	0,5
	Group 7	3,2
	Group 8	0,2
	Group 9	4,5
	No reply	15,3
Years of labour experience	1-20	73,3
	21-41	22,1
	No reply	4,5
Math capacity	Very good	10,3
	Good	36,2
	Sufice	44,1
	Weak	8,2
	No reply	1,1

Regarding the education levels of the respondents (Table 2), those with bachelor's degrees (43.3%) and secondary education (24.4%) predominated. It was further observed that 15.3% of the respondents had education beyond a bachelor's degree. Only 11.5% completed

education through the 9th year and 4.0% through the 4th year, while 0.3% had no primary instruction.

The majority of the respondents (58.6%) had higher degrees, and of the respondents with degrees, only 13.9% reported that their degrees were in economics or the business sciences. With regard to the training of the respondents, the majority (62.4%) reported that they did not receive any training in the areas of economics or business sciences. In an analysis of the respondents' professions and a classification of the same, according to the National Classification of Professions [Classificação Nacional de Profissões (CNP)], it was found that the majority (37.8%) belonged to group 2, which corresponds to specialists in intellectual and scientific professions. Finally, the number of years of experience in the job market varied from 1 to 41 years, with 73.3% having between 1 and 20 years of experience. The questionnaire also asked how the respondents considered themselves as mathematics students, on a scale of very good/good/sufficient/weak. The majority responded sufficient (44.1%) and good (36.2%). These variables indicate the qualifications and capacities of the respondents. The information is shown in Table 14.

With regard to the work situation, the majority of the respondents were employees (72.1%) and only 11.0% were self-employed. Additionally, 10.2% of the respondents were unemployed, as shown in Table 3.

Table 3. Stratification of the collected sample relative to the professional situation (%)

Working situation	%
Work for other	72,1
Self-employed	11,0
Study and works	0,5
Unemployed	10,2
Retired	0,5
Study	1,0
Housekeeper	4,0
Other	0,3
No reply	0,5

With regard to the economic situation (Table 4), the income distribution was analysed, and 37.3% of the respondents were found to be in the 2001€ to 6000€ range, while 30.5% were in the 1001€ to 2000€ range. Overall, 19.4% of the respondents had a net household monthly income below 1000€. Only 1.1% of the respondents had a monthly income that exceeded 6001€. When asked about their household monthly financial situations, the majority of the respondents revealed that the situation was satisfactory (59.6%), while 19.7% reported that the situation was good/very good and 19.7% that the situation was bad/very bad.

Table 4. Stratification of the collected sample relative to the economic situation of the respondents (%)		
Variable		%
Income	<1.000€	19,4
	1.001€-2.000€	30,5
	2.001€-6.000€	37,3
	6.001€-10.000€	0,8
	>10.001€	0,3
	No replay	11,6
Financial situation	Very good	0,5
	Good	19,2
	Average	59,6
	Bad	17,6
	Very bad	2,1
	No reply	1,0

RESULTS

Level of economic literacy

The level of economic literacy is evaluated according to the percentage of correct responses. A summary of the statistics related to lato sensu economic literacy is shown in Table 5, with consideration of those that are more financially oriented. The questionnaire includes 29 questions, of which 22 concern economics and 7 concern finances.

Table 5. Percentage of correct responses (%)				
	Descriptive statistics			
	Average	S.d.	Min	Max
Overall	73,1	19,9	0,0	100,0

From a total of 29 questions that evaluate economic literacy and financial comprehension, the average number of correct responses was approximately 73.1%. On a scale of 0 to 20 (Scale of Portugal) the average score was approximately 14.5. This result demonstrates that the respondents had a good understanding of economic and financial subjects. The standard deviation was 19.9%, the minimum value was 0.0% and the maximum value is 100%.

Table 6. Percentage of correct responses to the 22 economic questions (%)				
	Statistics			
	Average	S.d.	Min	Max
22 'economic' questions	75,6	20,4	0,0	100,0
7 'financial' questions	63,8	23,6	0,0	100,0

Regarding economics only, the average result was 75.6% (Table 6), which is equivalent to approximately 17 correct answers out of 22. On the Scale of Portugal (0-20), the individuals had an average economic literacy level of 15.12. Ferreira (2010) assessed the economic knowledge of the same target population, and the average score obtained was 68.5%. Two years later the respondents improved their performance from an average result of 68.5% to 75.6%. Partial financial literacy is evaluated by the number of correct responses to the seven multiple-choice questions in group 3 of the questionnaire.

The average number of correct responses was 63.8%; thus, on average, all of the interviewees responded correctly to more than half of the seven questions about financial knowledge (4.47 questions). On the Scale of Portugal (0-20), the individuals showed an average financial comprehension level of 12.8.

Hence, overall, the average number of correct responses to economics questions (75.6%) was greater than the average number of correct responses to financial questions (63.8%). With the objective of analysing whether partial economic literacy is greater than partial financial literacy, the SPSS was used, as well as the non-parametric Wilcoxon Signed Ranks test (paired samples), because the data do not validate the assumption of normality. It was concluded that the partial economic literacy is significantly higher than the partial financial literacy at the usual levels of significance.

Table 7. Percentage of individuals with correct responses to the economic questions by subject (%)

	Statistics			
	Average	S.d.	Min	Max
Economics of consumer	90,1	17,9	0,0	100,0
Economis of producer	74,2	24,5	0,0	100,0
Economics finance	70,0	27,1	0,0	100,0
Role of government	65,9	31,0	0,0	100,0
International Economics	76,8	29,0	0,0	100,0

As explained above, the economic questions addressed the following areas: consumer economics, producer economics, financial economics, the economic role of the government and international economics. It can be observed that the respondents performed better in the area of "Consumer Economics", the question with the largest percentage of correct responses is also found in this area (question 6), as shown in Table 7. However, the economics area in which the respondents performed the worst was that related to "the economic role of the government", with an average of 65.9% correct responses, and the question in which the respondents performed the worst was also in this area (question 8). It is notable that, in the five addressed areas, a large disparity was confirmed between the individuals' knowledge within each area. Ferreira (2010)

also concluded that the area with the best respondent performances was consumer economics, and the area with the worst performances was government-related.

The existence of differences in the percentages of correct responses between economic areas was tested. Given the non-parametric nature of the data, the Friedman test was performed for repeated measurements at the usual significance levels (1%, 5% and 10%). Significant differences were verified between the economic areas.

The questionnaire included a set of questions related to the interest and importance attributed to the economy. During the inquiry, the individuals expressed their opinions relative to their interests in economic subjects in two of the questions. The first question evaluated whether the individuals followed economic-related subjects and news through the various means of communication. Overall, 44.9% reported that they frequently followed news about the economy. However, a considerable percentage of respondents (31.7%) mentioned that they rarely followed economic subjects and announcements.

The majority of the respondents demonstrated that they were reasonably interested in economic subjects, and a few respondents said that economic subjects were very interesting. This result was expected, given that individuals still do not recognise the importance that the economy has in their lives and in the world around them. This information is summarised in Table 8.

Table 8. Distribution of the obtained responses to questions related to the respondents' interests in economic subjects (%)

Question	**Answer**	**%**
How frequently you follow economic news? (magazines, newspapers, TV, radio or internet)?	Follow very frequently	18,4
	Follow frequently	**44,9**
	Follow rarely	31,7
	Do not follow	4,4
	Not reply	0,6
How would you rank your interest abou economic matters?	Very interested	19,5
	Reasonable interested	**64,6**
	Little interested	13,1
	Not interested	1,9
	Not reply	0,8

In addition to evaluating the respondents' interests in economics, it is also essential to perceive the importance of economic knowledge in financial and political situations and wealth. Thus, the respondents were questioned about the importance of economic knowledge in various situations. Table 9 indicates the degree of importance that the respondents placed on each one of the situations.

Table 9. Distribution of the obtained responses relative to the importance of economic knowledge in various situations (%)		
Question	Answer	**%**
How importante is economic knowledge for these following situations?		
To understand bettwe polititians promisses and actions	Very important	22,5
	Important	**46,0**
	Not much important	17,6
	Not important at all	11,1
To get a better job and better wage	Very important	20,7
	Important	**48,5**
	Not much important	21,8
	Not important at all	6,5
To be a better and more active citizen in society	Very important	28,4
	Important	**53,0**
	Not much important	11,8
	Not important at all	4,4
To take better decisions and manage better my investments and savings	Very important	**66,2**
	Important	26,7
	Not much important	2,4
	Not important at all	2,1
To take better decisions regarding present and future consumption	Very important	44,7
	Important	**45,2**
	Not much important	5,5
	Not importante at all	1,9
To manage better my debts	Very important	**59,9**
	Important	32,0
	Not much important	3,2
	Not important at all	2,4
To improve my wealth and wellbeing	Very important	31,3
	Important	**52,0**
	Not much important	10,7
	Not important at all	3,4

It was confirmed that the respondents thought that having economic knowledge was very important when making better decisions about investments and savings (66.2%) and also for the better management of decisions about loans and credit (59.9%). Thus, the respondents gave greater importance to the economy in financial situations. The majority of the respondents indicated that was is important to understand the economy in the remaining situations.

Because the respondents were parents/guardians and teachers of students in elementary education, it was considered pertinent to understand whether the respondents thought that the application of economic disciplines was relevant in basic education, as shown in Table 10.

When questioned about the importance of inserting economic subjects into basic education programs for students, the majority of the respondents (57.5%) considered it to be relevant. Only 25.5% considered it very relevant to include economic subjects in basic education programs, and 4.2% considered it irrelevant to educate the youngest students about economic science.

Table 10. Distribution of the respondents' responses to the question of the relevance of the insertion of economic subjects into Basic Education (%)

Question	Option	%
How importante you think it is to include economics into basic education?	Very relevant	25,5
	Relevant	**57,5**
	Not much relevant	11,5
	Not relevant at all	4,2
	Not answer	1,3

Determinants of economic and financial knowledge: Econometric Model and variables

A central aim of this paper is to explore factors that contribute to explain the performances of adults in terms of economic and financial literacy. To this end, a model was developed to consider a set of factors as explicative variables, which, according to the literature, may contribute to explanations of the differences in economic literacy between adults.

The multiple linear regression model was adopted as the econometric methodology; because this model only includes cross-sectional data (a sectional sample in which individual observations are obtained at the same moment in time), it establishes a relationship of dependence and has many exogenous variables (Wooldrige, 2006).

The model takes the following form:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + ... + \beta_k x_{ki} + \mu_i \quad (1)$$

The majority of the articles in the literature review used an ordinary least squares (OLS) estimate, and this was the option for the estimation in this study.

The dependent variable, y_i, EFTOTAL_PERC, corresponds to the percentage of the number of correct responses to the 29 questions, which varied from 0% to 100%. (following Walstad & Rebeck, 2002). We explore a number of explanatory variables (x_{ji}). AGE and AGE2 correspond, respectively, to the age of the respondent in years and the age of the respondent in

years squared. It is expected that economic literacy will increase with age, although at a decreasing rate.

GEN indicates the gender of the respondent. It is a dummy variable equal to 1 if the respondent is male and 0 if the respondent is female.

NAT corresponds to the nationality of the respondent and can be considered a proxy for ethnicity, which was significant in the study by Mandell & Klein (2007). It is a dummy variable that is equal to 1 if the respondent is Portuguese and 0 if they are of another nationality.

EDU1, EDU2 and EDU3 are variables related to the education level of the respondent. Thus, education level is aggregated into three main groups. EDU1 corresponds to the respondents that completed the mandatory education and is equal to 1 if they belong to this level and 0 if they do not. To avoid exclusion, the minority that responded that they did not have primary instruction or only completed up to the 4th year of education was also included in this group. EDU2 corresponds to the respondents that confirmed having secondary education and is equal to 1 if they belong to this level and 0 if they do not belong to this level. EDU3 indicates that the respondents had a degree (bachelors, masters or doctorate), and the value is equal to 1 if they belong to this level and 0 if they do not belong to this level. The category excluded for this set of dummy variables was EDU1, and its effect was captured in the constant term. It is expected that people with more education will perform better, as observed in the studies by Wood & Doyle (2002) and Walstad & Rebeck (2002).

ECON is a dummy variable that represents individuals with some type of degree in the areas of economics or finance. It is expected to assume positive values because the possession of this type of degree indicates that the individual will have more knowledge about the subject and consequently, better results.

NCP is an ordinal variable and comprises an evaluation of the respondents' professions according to the National Classification of Professions. Thus, this variable varies from 1 to 9, with 1 corresponding to the professions with the highest qualifications and 9 to the professions with the lowest qualifications. It is expected that the greater the qualifications of the profession, the better the obtained results will be. Thus, the value of this variable is expected to be negative because 1 is the highest qualification and 9 is the lowest qualification.

MAT is a dummy variable equal to 1 for individuals who consider themselves to be Very Good at mathematics, and 0 for the others. It is expected that the individuals with more mathematics knowledge will have better results.

ACTIVE is a dummy variable and takes the value of 1 in cases of employed respondents (encompasses both employees and self employed respondents) and 0 for non-employed respondents (e.g., unemployed, retired, home maker, student, other). It is expected that the active respondents will have better results than those that are not active.

NMI corresponds to the household net monthly income of the respondent. This variable is continuous and varies from 1 to 5. A value of 1 corresponds to lower income levels, and 5

corresponds to the highest income levels. It is expected that this variable will positively affect the respondents' performance.

Table 11. Main variables of the study		
Variable		**Decription**
Dependent variable		
Percentage of correct answers (EFTOTAL_PERC) Percentage of correct answers, varies 0 to 100%		
Explanatory		
Age (AGE and AGE2)		Age in years and Age in years squared
Gender (GEN)		1=male; 0=female
Nacionality (NAC)		1=portuguese; 0= other
Education level	ESC1	Up to 9 years Education : 1= Yes; 0= no
	ESC2	Secundary: : 1= Yes; 0= no
	ESC3	University degree : 1= Yes; 0= no
Add training in economics / finance (ECON)		1= Yes; 0= no
Professional qualification level	CNP	1 to 9
Math capacity (MAT)		Good in maths: : 1= Yes; 0= no
Working (ACTIVE)		Is working: : 1= Yes; 0= no
Net monthly income (NMI)		1 to 5

The list of variables is summarised in Table 11.

Econometric results

The estimates for the number of correct responses are shown in Table 9. In the OLS regression, the 43.03% variation in global literacy was explained by the variables of the model. The coefficient obtained from AGE and AGE2 reflectes an inverted U relationship. The gender variable (GEN) was significant and positive. Thus, males have global literacy levels that are higher relative to the females, if the other explicative variables are constant. This result goes is in line with the studies by Walstad & Rebeck (2002) and Wood & Doyle (2002) for other economies.

Nationality (NAT) is a significant variable, as it was in the study by Mandell & Klein (2007), and Portuguese individuals were found to perform better than individuals of other nationalities, ceteris paribus. This result can be explained by differences in native language because non-Portuguese people might not be familiar with the economic terms or the Portuguese economic realities.

Relative to education, those who had a secondary education (EDU2) had more percentage of correct responses than those who had a basic education, ceteris paribus. When the respondents with higher education (EDU3) were analysed, the average score was even higher relative to those with only a basic education, ceteris paribus. Thus, it is confirmed that the effects of the education level are more evident and positive for those respondents with degrees, although both were significant. This result agrees with those obtained by Wood & Doyle (2002) and Monticone (2010).

Previous training in business sciences (ECON) also influenced the global literacy of the respondents because respondents with training had better results than those who did not have training in economics, ceteris paribus. Walstad & Rebeck (1999) also concluded that training in economics had a positive effect on the level of economic literacy. This result, despite being positive, was very small and could be explained by the fact that the test was short and did not precisely measure what people learned during economic training, the dissipation over time of the economic knowledge gained in the training, variability in the quality of economics teachers and the materials used to teach economics, which could reduce teaching effectiveness, and finally the effects of guessing on a multiple-choice test, which could influence the scores of those who lacked economic knowledge.

Not surprisingly, the national classification of professions (NCP) was found to have a negative and significant impact on global literacy. This indicates that the respondents with weaker professional qualifications answered fewer questions correctly (3%) than did those with better professional qualifications, ceteris paribus. This can be explained by the fact that professions with better qualifications require that the respondents have a higher level of education and are more informed about societal problems while professions such as those in group 9 do not require a high level of education and consequently the respondents have less knowledge about economics and finances.

Another variables positive related to the economic literacy level is the level of maths of the individual. Individuals who considered themselves better at maths had higher scores than those who considered themselves worst in maths.

Individuals who were active had higher scores than those who were were not currently working, ceteris paribus. This variable indicates, not surprisingly, that working respondents had better economic knowledge. This result indicates that adults obtain economic information through various sources, such as friends, relatives and work colleagues.

In turn, the estimated coefficient for income is also statistically significant, indicating that, on average individuals from households with higher levels of net monthly income also have a higher economic knowledge, with everything else constant.

Table 12. Determinants of economic literacy		
	Estimated coefficients (t statistics)	
C	-7.766	(-0.334)
AGE	2.269**	(2.104)
AGE2	-0.025**	(-2.003)
GEN	4.656*	(3.443)
NAC	14.239*	(4.400)
ESC2	7.508*	(2.715)
ESC3	8.344*	(2.702)
ECON	5.494*	(4.076)
CNP	-1.941*	(-4.178)
MAT	4.515**	(2.395)
ACTIVE	7.404*	(2.781)
NMI	4.320*	(4.245)
N	422	
R2	0.43025	
R2-ajustado	0.41496	
LR statistic	28.1464	
Prob(LR statistic)	0.00000	

DISCUSSION

This study fills an empirical investigation gap and calls attention to a question of extreme interest, namely the economic literacy of a general population. This paper is part of an investigation project of the Department of Economics, Engineering and Industrial Management of the University of Aveiro, Economicando, which is financed by the Foundation for Science and Technology (Fundação para a Ciência e a Tecnologia - FCT).

The consensus seems to be that economic literacy is increasingly important, given the growing complexity and variety of financial products and services available on the market, as well as the perceptions of the conditions and realities in which this set of economic activities have developed. Despite the growing attention paid to the dissemination of economic science, empirical studies show that individuals have little knowledge of economics and finances, and thus it is necessary to define policies that will increase individual interest in the knowledge of economic subjects. Based on the literature review, factors were identified that permit an explanation for the levels of economic and financial literacy in the general population, as well as why the levels differ. Out of a total of 29 questions that evaluated economic literacy and financial comprehension, the average number of correct responses was 21 (73.1%), which translates to a good level of economic knowledge on the part of the respondents. However, when comparing the results from the economic and financial questions, it was found that individuals

performed better on the economics questions. The respondents were in general interested in the subject and considered it important to have economic knowledge in various situations, mainly financial situations.

Various international studies have explored ways to improve the economic literacy levels. Monticone (2010) and Huston (2010) reported that active measures were needed to create a financially responsible work force. More education, dissemination of information, transparency of financial institutions and greater access to financial counselling are necessary, especially for the most vulnerable individuals. In a more comprehensive manner, the government could contribute to improved economic literacy in the general public by promoting the integration of economic subjects in all schools and means of communication (Federal Reserve Bank of Minneapolis, 1999). To this end, it will be necessary to train teachers to increase their economic knowledge and develop their manners of thinking about economic subjects.

With the OLS model we tested the importance of a number of explanatory variables.

For future research it would be interesting to implement the same evaluation tool to a representative sample of all Portuguese population at another period of time and confirm the evolution of respondent knowledge. Another suggestion for further investigation involves a study of the effects of economic literacy on the attitudes and well being of the individual. In this study, factors that affect economic literacy were studied, but economic literacy is thought to affect other variables, thus making it an explicative variable.

It is expected that this work will contribute to an increased interest in "education in economics" on the part of researchers and that their results will allow for the expansion of knowledge about the Portuguese reality, being possible to compare the results to others obtained internationally. The developed questionnaire can also be applied by other researchers in the future.

*AKNOWLEDGEMENTS

The study has been conducted under Research project "Economicando" (PTDC/EGE-ECO/100923/2008), financed by FEDER funds through the Programa Operacional Fatores de Competitividade - COMPETE and by national funds through the FCT - Fundação para a Ciência e Tecnologia.

REFERENCES

Banco de Portugal. (2011). Relatório do Inquérito à Literacia Financeira da População Portuguesa: 2010. Banco de Portugal, Eurosistema, Lisbo,. from: www.bportugal.pt

Belbute, J.M.M., & Sousa, M. (2004). Literacia Económica entre alunos Universitários, Documento de Trabalho nº 2004/06, Universidade de Évora, Departamento de Economia.

Clark, J.R., Schug, M.C., & Harrison, A.S. (2009). Recent Trends and New Evidence in Economics and Finance Education. *Journal of Economics and Finance Education*, 8(1), 1-10.

Federal Reserve Bank of Minneapolis. (1999). Economic Literacy Action Steps. *The Region, Federal Reserve Bank of Minneapolis*, 13(2), 8.

Ferreira, A. (2010). *Um Contributo para a Divulgação da Ciência Económica*. Departamento de Economia, Gestão e Engenharia Industrial, Universidade de Aveiro.

Gleason, J., & Van Scyoc, L. J. (1995). A Report on the Economic Literacy of Adults. *Journal Of Economic Education*, 26(3), 203-210.

Grimes, P., Millea, M., & Thomas, M. (2010). Testing the Economic Literacy of K-12 Teachers: A State- Wide Baseline Analysis. *American Secondary Education*, 38(3), 4-20.

Gujarati, D. (2004). *Basic Econometrics,4th Edition.* New York: The McGraw-Hill Company.

Huston, S.J. (2010). Measuring Financial Literacy. *The Journal of Consumer Affairs*, 44(2), 296-316.

Jappelli,T. (2010). Economic Literacy: An International Comparison. *The Economic Journal, Royal Economic Society*, 120(548), 429-451. doi: 10.1111/j.1468-0297.2010.02397.x.

Koshal, R., Gupta, A., Goyal, A., & Choudhary, V. (2008). Assessing Economic Literacy of Indian MBA Students. *American Journal Of Business*, 23(2), 43-51.

Monticone, C. (2010). How Much Does Wealth Matter in the Acquisition of Financial Literacy?. *Journal Of Consumer Affairs*, 44(2), 403-422. doi:10.1111/j.1745-6606.2010.01175.x

Nunnaly, J.C. (1978). *Psychometric theory*. New York: McGraw-Hill Inc

Pang, M.F. (2010). Boosting financial literacy: benefits from learning study. *Instructional Science*, 38(6), 659-677. doi:10.1007/s11251-009-9094-9.

Remund, D.L. (2010). Financial Literacy Explicated: The Case for a Clear Definition in an Increasingly Complex Economy. *The Journal of Consumer Affairs*, 44(2), 276-295.

Schuhmann, P., McGoldrick, K., & Burrus, R. (2005). Student Quantitative Literacy: Importance, Measurement, and Correlation with Economic Literacy. *American Economist*, 49(1), 49-65.

Shipley, C., & Shetty, S. (2008). Factors Affecting Economic Literacy of College Students: some additional evidence. *Journal of Economics and Economic Education Research*, 9(1), 13-18.

Tabesh, H., & Schultz, B. L. (2007).Determinants of High School Economic Literacy: A Case Study. International *Journal Of Business Research*, 7(6), 28-34.

Walstad, W. B., & Rebeck, K. (2002). Assessing the economic knowledge and economic opinions of adults. *Quarterly Review of Economics and Finance*, 42(5), 921-935.

Walstad, W. B., Rebeck, K., & MacDonald, R. A. (2010). The Effects of Financial Education on the Financial Knowledge of High School Students. *Journal Of Consumer Affairs*, 44(2), 336-357.

Walstad, W., & Rebeck, K. (1999). How Does Economic Education Impact Economic Literacy?. *Region* (10453369), 13(2), 18.

Walstad, W.B.,& Robson, D. (1990). *Basic Economics Test: Examiner's Manual. Joint Council on Economic Education*, New York.

Walstad, W.B.,& Soper, J.C. (1987). *Test of Economic Knowledge: Examiner's Manual. Joint Council on Economic Education*, New York.

Wood, W. C., & Doyle, J. M. (2002). Economic Literacy among Corporate Employees. *Journal Of Economic Education*, 33(3), 195-205.

Wooldrige, J. (2006). *Introdução à Econometria: uma abordagem moderna*, Thomson, São Paulo.

DETERMINANTS OF COLLEGE BASKETBALL GRADUATION RATES

Neil Terry, West Texas A&M University
Anne Macy, West Texas A&M University
John Cooley, West Texas A&M University
Ashley Peterson, West Texas A&M University

ABSTRACT

This paper presents empirical results investigating the determinants of six-year graduation rate for college basketball teams. The research sample consists of 434 (217 men's and 217 women's) college basketball programs during the years 2004-2010. Demographic and performance data are from the 2008 college basketball season. Significant positive determinants of college basketball six-year graduation rates are profitability of the overall athletic program, size of the institution defined by number of undergraduate students, financial support the institution offers to college athletes, recruiting budget of the athletic program, percent of the recruiting budget allocated to female athletes, number of team wins, and categorical trait of being a women's team over a men's. The empirical results indicate classification as a public institution and percent of the financial support allocated to female athletes at an institution are negative and have a statistically significant impact on six-year graduation rates of basketball programs. Profitability of the basketball program and average pay for head coaches are not statistically significant determinants.

INTRODUCTION

Universities receive their non-profit status thanks to their role of educating students but the business aspect of college sports continues to grow and expand. The University of Texas leads collegiate athletic programs with over $120 million a year in revenue generation, which include approximately $15 million generated by the men's basketball program and $3 million by the women's basketball program. Critics of college sports cite the revenue generated by athletics as evidence of their commercial nature. Supporters counter by stating the overall goal of athletics is not to turn a profit but to provide financial support to student athletes and increase the university's national profile (McEvoy, 2005; Smith, 2008). Proponents of major college athletics highlight the positive externalities associated with the public relations and institutional branding produced by successful athletic programs (Smith, 2008). The role of athletics on a college campus can be debated but graduation rates measuring the proportion of an entering class that have graduated within a specific number of years are one of the most common outcome

measures. Scott, Bailey, and Keinzl (2006) argue for using six-year graduation rates as a performance measure because it is one of the most important measures, is a measure available for a large number of institutions, and allows comparable findings to other results in the literature.

Athletics is a driving force at many institutions of higher education. The purpose of this research is to investigate the determinants of six-year graduation rates for college basketball programs. The determinants model considers multiple variables including athletic program profits, basketball program profits, winning, institution size, recruiting budget, men's versus women's program, public versus private institutions, and financial support. The organization of the manuscript is as follows: The first section offers a brief review of the literature. The second section describes the data and model. The next section offers empirical results for the determinants of six-year graduation rates for college basketball derived from 434 college basketball programs. The final section offers a summary and conclusions.

SURVEY OF THE LITERATURE

One of the most pressing issues facing American universities is the number of students who fail to graduate. Low graduation rates cost universities scarce resources; weaken the ability to meet educational objectives; and are perceived to reflect the university's ability to meet the educational, social, and emotional needs of students (Mangold, Bean & Adams, 2003). There is a dearth of research on the graduation rates of college athletes and athletic programs but there is an established independent research track for both graduation rates and various aspects of college athletics. Retention rate has dominated studies looking at academic persistence. Academic and social attachment currently forms the foundation of most research on persistence and graduation success (Pasarella & Terenzini, 1991; Tinto, 1993). Institutional or social policy designed to increase retention generally focus on strengthening student attachment, for example through improving student services or increasing intramural and varsity athletics. Metzger and Bean (1987) find that age and goals have a greater role in persistence and related outcomes for non-traditional than traditional students.

Mangold, Bean, and Adams (2003) find a negative relationship between athletic success and graduation rates at NCAA Division I institutions. Successful intercollegiate sports may not provide a mechanism for academic integration and may, under certain conditions, actually weaken it. In order to resolve this possible conflict between the results and the existing literature, the authors begin by pointing out that social involvement, if carried too far, can result in suboptimal outcomes. Many of the factors that inhibit social integration may also weaken academic integration and attenuate persistence (such as commuting, maintaining friendships with peers not attending college, off-campus employment). In addition, activities that are not part of the student's academic environment, such as commuting or off-campus employment, may also weaken academic and/or social integration and thus compete with learning objectives as well as a student's overall commitment to graduation. Their results suggest that social involvement in

intercollegiate sports, a process that broadly and indirectly is expected to facilitate graduation, may work in combination with other institutional characteristics to inhibit it.

The student demographic characteristics are often different for public and private schools. Public institutions tend to have relatively larger numbers of commuter and older students. Scott, Bailey, and Kienzl (2006) employ a selectivity measure via high school GPA or SAT admission scores as a proxy for quality. Private institutions tend to have higher admission traits than public institutions. Scott, Bailey, and Kienzl (2006) show that public colleges have lower six-year graduation rates than private colleges but if resources and student populations are controlled, public colleges are able to do more with less and graduate a slightly larger percentage of students. Astin and Oseguera (2002) employ regression analysis for their empirical work, which reveals institution type (private, public, college, university), SAT score, GPA, race, and gender all have an impact on retention and graduation rates. Importantly, they find that the gap in six-year graduation rates between public and private colleges diminishes significantly, from 31% to about 7% when all these factors are controlled.

Rishe (2003) uses least squares estimates from Division I schools to examine how athletic success influences graduation rates. He finds that neither the graduation rate for student-athletes nor graduation rate for all other undergraduates is sensitive to the level of a school's athletic success. However, the graduation gap between student-athletes and all other undergraduates is sensitive to various measures of a school's athletic success. Women have higher graduation rates than men in general, and this gender graduation gap is exacerbated when focusing on student-athletes at schools with the most prominent football programs.

The success of collegiate athletic programs might have an indirect impact on an academic institution or the local community. Tucker (2005) finds a statistically significant impact for successful football teams on the quality of incoming freshman class, which provides evidence of a strong athletic advertising effect for football. Multiple studies find a positive correlation between athletic success and alumni giving rates (Rhoads & Gerking, 2000; Turner, Meserve & Bowen, 2001; Monks, 2003; Holmes, Meditz & Sommers, 2008). Rees and Schnepel (2009) find host communities register sharp increases for assaults, vandalism, and arrest for disorderly conduct on college football game days. Upsets are associated with the largest increase in the number of expected offenses. Baade, Baumann, and Matheson (2008) examine the economic impact of college football on the local economy. The research focuses on 63 metropolitan areas that played host to major college football with a research sample from 1970 through 2004. Number of home games played, winning percentage of local team, and winning a national championship are shown to have an insignificant impact on employment and personal income in the cities where the teams play. Lentz and Laband (2009) examine the economic impact of college athletics on employment in the restaurant and accommodations industries. They find a positive and statistically significant relationship between college athletics revenue and MSA employment in the food services and accommodations industries. Siegfried, Sanderson, and

McHenry (2007) argue that the economic impact analyses developed by most college and universities tend to inflate the real economic impact.

Terry, Pjesky, and Patterson (2011) investigate the determinants of men's college basketball profit. Size of the student athlete recruiting budget, size of the institution measured by number of undergraduate students, availability of financial aid to student athletes, head coach compensation, and winning are revealed to be positive and statistically significant determinants of men's college basketball profitability. Model results imply profitability of the overall athletic program at an institution, percent of student athletic financial support allocated to women, public institutions, number of female athletes at an institution, and compensation for assistance coaches are not significant determinants of men's college basketball profits.

Compensation of college coaches can have a significant impact on the performance of an athletic program. Terry, Pjesky, and Rider (2009) conclude the significant determinants of head coaches pay are profitability of the athletic program, recruitment budget, percentage of the recruitment budget allocated to women's sports, compensation of assistant coaches, number of female athletes at the institution, and number of sports supported by the athletic program. The Equal Employment Opportunity Commission (EEOC) has ruled all collegiate coaching jobs are substantially equal. All coaches at all levels perform certain functions including teaching/training, counseling/advising student athletes, general program management, budget management, fundraising, public relations, and recruiting at the college level. Labor market theories suggest similar individuals who do the same job with the same support should earn similar salaries. Brown and Jepsen (2009) find this to be true among major league baseball players. Players with higher offensive statistics (productivity) did receive higher salaries. Idson and Kahane (2000) find that having productive teammates enhances productivity and compensation. Kahn (2006) found that African-American coaches were not victims of discrimination in wage, hiring, or firing in the NBA. Humphreys (2000) reports that male head coaches of women's basketball teams earn less than do female head coaches of women's basketball teams.

Title IX prohibits any type of gender discrimination in any educational programs or activities within an institution receiving federal financial assistance. The act applies to both public and private schools, from kindergarten through graduate school, and covers admission, recruitment, educational programs and activities, course offerings and access, counseling, financial aid, employment assistance, facilities and housing, health and insurance benefits and services, scholarships, and athletics. Title IX has been the most important measure ever undertaken to promote gender equality in sports (Leeds & Von Allen, 2002). From 1971-2002 the number of women in college sports increased fivefold. In fact, now there are more women's teams than men's teams: 9,479 to 9,149. The potential conflict with the expansion of women's athletics is the redistribution of football profits to female non-revenue generating sports at the expense of male non-revenue generating sports like wrestling and rugby (Terry & Ramirez, 2005).

The literature implies size of a college via number of students could have a positive

impact on athletic program profitability. The labor economics literature has revealed the tendency for large firms to be more profitable and pay employees more than small firms (Lucas, 1978; Oi, 1983; Brown & Medoff, 1989; Fox, 2009). Absolute profits and profit rates both have a tendency toward positive correlation with size. Large state universities like the University of Texas, University of Michigan, and University of Florida might have an innate advantage with respect to football programs based on their dominant size. The extra profits might indirectly influence six-year graduation rates by offering large institutions a larger resources base for student support that helps facilitate graduation.

DATA AND MODEL

The NCAA regularly surveys member institutions to assess compliance with Title IX and other regulations. This study uses data from the 2008 Office of Postsecondary Education Equity in Athletics Disclosure website and 2004-2010 six-year graduation rates from the NCAA Graduation Success Rate website. The research cohort is derived from 434 Division I (217 men's and 217 women's) Division I collegiate basketball programs. The explicit empirical model employed to investigate the graduation rate for college basketball is specified as follows:

$$(1)\ GR_i = B_0 + B_1APROFIT_i + B_2PPROFIT_i + B_3STUDENTS_i + B_4PUBLIC_i + B_5FINSUPP_i + B_6WFINSUPP_i + B_7RECRUIT_i + B_8FRECRUIT_i + B_9COACH_i + B_{10}WINS_i + B_{11}FEMALE_i + u_i.$$

Table 1 presents summary statistics for model variables. The dependent variable GR measures six-year graduation rates for basketball programs for the years 2004-2010. A six-year graduation rate was selected in order to be consistent with the literature and employ a measure that is more flexible to life challenges than the timely four-year graduation rate. Seventy-nine college basketball programs reported a six-year graduation rate higher than 95%. Fifty-eight of the seventy-nine programs reporting 95% or higher graduation rate represent women's college basketball. Nine institutions in the research cohort have both men's and women's college basketball six-year graduation rates higher than 95%. The nine programs are Yale University, Brown University, Harvard University, Dartmouth College, Princeton University, Bucknell University, University of Dayton, University of Rhode Island, and Villanova University.

The model includes eleven independent variables. Two of the variables focus on profitability. The expectation is for the profit variables to have a positive impact on six-year graduation rates based on the expectation that profits have a positive impact on resources, which include the luxury of smaller class sizes and the availability of tutors. The variable APROFIT controls for the profit of the overall athletic program at an institution. Notre Dame ($26.1 million), University of Michigan ($20.8 million), University of Texas ($15.7 million), and University of Florida ($15.6 million) are the four athletic programs in the sample reporting the highest profitability across the entire athletic program. The sample cohort includes 204

institutions reporting overall athletic profits of less than $100, although no institution in the sample reports a negative overall profit for the athletic program. The variable PPROFIT measures basketball program profit (reported basketball program revenue minus basketball program cost) at the institution. Twenty-two programs in the research cohort report a basketball program profit of $5 million or higher. The three programs with the highest profitability are the men's programs at University of North Carolina ($11.6 million), University of Arizona ($13.2 million), and University of Louisville ($17.1 million). On the other hand, not all college football programs are profitable. In fact, four institutions (University of Akron, University of Tulsa, Villanova, and Ball State) report losses in excess of $3 million. In contrast, 152 of the women's basketball programs earned a negative profit.

Table 1
Summary Statistics

Variable	Mean	Maximum	Minimum	Std. Deviation
GR	0.7655	1.00	.10	0.1971
APROFIT	3,590,884	13,225,139	128,952	2,848,943
PPROFIT	204,707	17,134,624	-3,378,575	2,175,283
STUDENTS	12,937	36,612	1,678	8,223
PUBLIC	0.74	1	0	0.4381
FINSUPP	4,753,853	15,478,248	0	2,841,703
WFINSUPP	0.4152	0.6100	0	0.0976
RECRUIT	494,329	2,005,677	28,500	366,236
FRECRUIT	0.3281	0.6461	0.0941	0.0723
COACH	192,864	903,890	16,674	170,443
WINS	15.38	30	2	6.33
FEMALE	0.50	1	0	0.50

The independent variables STUDENTS and PUBLIC are institutional control variables. The STUDENTS variable captures the size of the institution. STUDENTS is a measure of the number of undergraduate students enrolled at the institution. The largest institution in the sample is Penn State with 36,612 undergraduate students, while the smallest institution in the research cohort is Davidson College with 1,678 students. The impact of institution size on six-year graduation rates could be positive or negative. The positive attribute is that large institutions can take advantage of economies of scale with respect to providing student support services. The negative attribute for large institutions is the prospect of larger class sizes and less personal attention per student. The variable PUBLIC is a categorical variable controlling for public versus private institutions. Public institutions represent seventy-four percent of the institutions in the research sample. The expectation is for public institutions to have a lower six-year graduation rate than their private counterparts based on private institutions ability to recruit student athletes with stronger academic backgrounds.

The next four independent variables in the model are resource control variables. FINSUPP is the amount of financial aid support available to students at an institution. Stanford ($15,478,248) and University of Notre Dame ($13,793,174) are the two institutions that offer the greatest amount of financial support to student athletes. In contrast, several of the Ivy League programs, including Yale, Dartmouth, and Harvard, do not explicitly offer financial support to students based on their classification as a student athlete. Despite the Ivy League programs not explicitly providing support to student athletes, the expectation is for financial support to have a positive impact on graduation rates. The variable WFINSUPP measures the percent of financial support in the athletic department allocated to female athletes. Drake University leads the research cohort in percentage of support allocated to women at 61%. A higher allocation of financial support to female athletes should have a negative impact on six-year graduation rates for men's programs but a positive impact on women's programs. RECRUIT is the budget allocated to the athletic department to recruit student athletes. The largest recruiting budget in the sample is $2,005,667 at the University of Tennessee, while the smallest reported recruitment budget is $28,500 at Texas Southern University. Recruiting budget should have a positive impact on the graduation rate of basketball programs assuming higher recruiting budgets offer programs the ability to attract individuals with both athletic and academic acumen. The variable FRECRUIT measures the percent of the recruiting budget in the athletic department allocated to female athletes. South Carolina State University leads the research cohort in percentage of recruiting budget allocated to women at 65%. A greater percentage of recruitment funds allocated to female athletes should have a negative impact on the six-year men's basketball graduation rate but enhance the women's basketball graduation rate.

The variable COACH is defined as the average pay of head coaches in male or female sports at the institution. The COACH variable serves as a proxy for compensation of head coaches for the men's and women's basketball programs, which should be highly correlated with average head coach pay at an institution. The University of Texas and University of Kansas lead the way with average head coach salaries of $903,890 and $748,953, respectively. Saint Peters College offers the lowest average head coach salary at $16,674.

The variable WINS measures the number of basketball wins for the 2008 season. Three men's basketball programs won 30 games in 2008 (University of Kansas, University of Memphis, and Ohio State University). Six women's basketball programs won at least 28 games in the research sample year (University of Tennessee, Stanford University, University of Connecticut, University of North Carolina, Purdue University, and Bowling Green University). The women's program at Iona College and the men's program at California State University at Sacramento are at the bottom of the winning list with only two wins in the season ending in 2008. Winning might have a positive impact on six-year graduation rates if winning increases student engagement into campus life and serves as a motivator to maintain athletic eligibility via academic performance.

The final variable is a categorical variable separating women's basketball programs from men's basketball programs. The average graduation rate in the sample is 76.5% but a cursory look at the numbers provides evidence that women's programs perform at a significantly higher rate than the men's programs. Specifically, the average six-year graduation rate for men's basketball programs is 67.7% while the six-year graduation rate for women's basketball program is 85.4%.

RESULTS

Table 2 presents the estimated empirical relationship between the explanatory variables and six-year graduation rates of college basketball programs. The ordinary least squares (OLS) model explains over 54 percent of the variance in college basketball six-year graduation rates. A model with logarithmic transformations of the dependent variable was considered but was not substantially different from the parsimonious OLS model. The alternate specification raised the R-square to over 56 percent but did not fundamentally change the significance or relative magnitude of any of the independent variables. None of the independent variables in the model have a correlation higher than 0.6, suggesting that excessive multicollinearity is not a problem in the analysis. Nine of the eleven variables in the model are statistically significant.

Table 2
Estimation of Equation 1: Determinants of College Basketball Graduation Rates (2004-2010)

Variable	Coefficient	t-statistic
Intercept	0.720009	13.26
APROFIT	1.181E-08	2.35*
PPROFIT	-1.719E-10	-0.04
STUDENTS	2.728E-07	2.06*
PUBLIC	-0.156036	-6.97*
FINSUPP	8.485E-09	2.24*
WFINSUP	-0.170689	-1.93*
RECRUIT	1.416E-07	3.59*
FRECRUIT	0.200062	1.99*
COACH	-1.507E-06	-1.32
WINS	0.004059	3.10*
FEMALE	0.138061	7.24*

Notes: R-square = .4862, F = 23.86, *p<.05, and n = 96.

The first two variables in the model are APROFIT and PPROFIT, which measure the impact of different measures of profitability on the graduation rates of college basketball teams, holding other variables constant. The APROFIT variable is positive and statistically significant. Clearly, a profitable athletic program has a positive and significant relationship with the graduation rates of college basketball program. Athletic programs that earn relatively high

profits tend to have more resources that can support the academic success of student athletes, such as tutors and other forms of individual instruction. The PPROFIT variable is negative but not statistically significant. The negative coefficient associated with basketball program profit is surprising given the positive coefficient associated with athletic program profit. One possible explanation for the negative impact of basketball profits on graduation rates might be that the goal of players on high-profile college teams that earn big profits is not to graduate but to facilitate a professional career. Most collegiate basketball players probably have some degree of professional aspirations but athletes playing for elite teams that are highly profitable have a more realistic aspiration. In addition, over half of the players drafted in the National Basketball Association (NBA) are usually underclassman that did not complete degree requirements. It is also possible that some men's college basketball programs that earn high profits will view placement in the NBA as a more important goal than graduating athletes, a sentiment that has been echoed by University of Kentucky Coach John Calipari in recent years. Placement in the NBA gives a program a long-run recruiting and sustainability advantage. Programs can recruit top high school talent with the hook that the college basketball program is a factory that produces professional athletes earning millions in the NBA.

Both of the institutional variables are statistically significant. Size of the institution measured by number of undergraduate students (STUDENTS) is a positive and statistically significant determinant of six-year graduation rates of basketball programs. The economies of scale of the resource base at a large institution might help students with tools for academic success. Possible advantages at large institutions include additional tutor support, personal mentorship, and other student support services that help persistence and graduation rates. The PUBLIC variable has a negative coefficient that is highly significant. The regression coefficient indicates that public institutions have a six-year basketball graduation rate that is approximately 15.6% lower than private institutions. The admission standards at private institutions are often higher than standards at public institutions. The higher standards might hurt the ability of private institutions to attract athletes with marginal academic ability but should help attract athletes with greater academic ability. Private institutions with strong academic reputations like Duke University are able to attract elite talent (e.g., Danny Ferry, Christian Laettner, Bobby Hurley, Grant Hill, Elton Brand, Jay Williams, Shane Battier, Mike Dunleavy, Luol Deng, Shelden Williams, Kyrie Irving, and Austin Rivers are all Duke players selected within the first ten picks in an NBA draft) but admission to Duke requires demonstrated academic acumen that is not dismissed by athletic prowess. With the historical exception of Duke University (1991, 1992, 2001, and 2010), Georgetown (1984), and Marquette (1977), private institutions and their higher academic admission standards have fallen short of national championships in the last fifty years. Higher academic standards might limit recruiting opportunities for private institutions but appear to have a positive influence on the six-year graduation rates of basketball programs.

All four of the resource control variables employed in the empirical model are statistically significant. FINSUPP is the amount of financial aid support available to students at

an institution. Not surprisingly, the FINSUPP regression coefficient is positive and statistically significant. Institutions that can afford to offer more financial support have the resources to help students achieve graduation success. The variable WFINSUPP measures the percent of financial support in the athletic department allocated to female athletes. The empirical results verify the hypothesis that a higher allocation of financial support to female athletes has a negative impact on six-year graduation rates for college basketball programs. Consistent with the findings of Rishe (2003), the results imply that male athletes in prominent athletic programs like basketball need more relative support if there is a desire to close the gender graduation gap. Athletic programs with a relatively large recruiting budget (RECRUIT) appear to find more success with respect to graduating basketball players. The RECRUIT variable has a positive and statistically significant coefficient. Higher recruiting budgets offer programs the ability to attract individuals with both athletic and academic acumen. It is almost certain that winning is a primary goal for most college basketball programs but coaches and recruiters also prefer to run a program that graduates student athletes because higher graduation rates creates positive externalities with university administrators and can be a recruiting tool with parents. The variable FRECRUIT measures the percent of the recruiting budget in the athletic department allocated to female athletes. The FRECRUIT variable has a positive and statistically significant impact on six-year basketball graduation rates. A possible explanation is that an athletic program that is cognizant of Title IX issues and makes an aggressive effort to support women's athletics is an athletic program that is also going to push for relatively high six-year graduation rates for all athletic programs. Programs that spend relatively more recruiting female athletes might have goals that are more process, equity, and academic outcome driven than simply having a winning men's football or basketball team that is a cash cow.

The head coach pay (COACH) variable has a negative but statistically insignificant impact on six-year graduation rates of college basketball programs. Head coaches receive compensation to do a variety of things for an athletic program but the empirical results of this study clearly indicate graduating student athletes is not one of the responsibilities. In fact, graduating players has a negative impact on head coach compensation. College athletics is enormously popular in the United States, and there is evidence that its appeal is growing (Rees & Schnepel, 2009). Winning games and energizing the alumni base is probably a more important determinant of the pay of college coaches than graduating students, although it is important to note this inference is limited by the observation the model employs a proxy for head coach pay.

The next variable in the model investigates the impact of winning (WINS) basketball games on six-year football graduation rates. Winning programs are more likely to help student athletes remain academically eligible for competition, which indirectly helps student athletes make positive progress toward degree completion. Winning might also have a positive impact on student athlete engagement into academic life and this engagement can augment retention and graduation rates (Scott, Bailey & Kienzl, 2006). Evidence from this cohort provides statistical evidence that winning has a positive and significant impact on six-year basketball graduation rates.

The final variable model is a categorical variable controlling for gender. The results indicate that women's basketball programs have a significantly higher graduation rate than men's basketball teams. Holding other variable constant, the women's graduation rate is approximately 13.8% higher than the men's six-year college basketball graduation rate. Higher female graduation rates are part of a national trend. At public institutions, about 58 percent of females seeking a bachelor's degree graduated within 6 years, compared with 53 percent of males; at private nonprofit institutions, 67 percent of females graduated within 6 years, compared with 63 percent of males (U.S. Department of Education, National Center for Education Statistics, 2012). Hence, the average six-year graduation rate for men's basketball programs at 67.7% and women's basketball programs at 85.4% are both relatively high. That being said, the extremely high graduation rate for women's college basketball programs provides evidence of the positive impact participation in athletics can be for women as part of an effort to achieve the goal of earning an college degree.

CONCLUSION

This study investigates the determinants of six-year graduation rates for college basketball programs. The research sample consists of 434 (217 men's and 217 women's) college basketball programs during the years 2004-2010. Profitability of the overall athletic program, financial support the institution offers to college athletics, recruiting budget of the athletic program, size of the institution defined as number of undergraduate students, and percent of the recruiting budget allocated to female athletes are revealed to be positive and statistically significant determinants of college basketball six-year graduation rates. The positive and significant variables lead to an overall conclusion that financial resources via profits, financial support to students, resources for support services provided by large institutions, and recruiting budget are keys to successfully graduating college basketball players. Winning games and the female categorical variable are two other variables that have a positive and statistically significant impact on college basketball graduation rates. It is interesting to note that, holding other variable constant, the women's graduation rate is approximately 13.8% higher than the men's six-year college basketball graduation rate.

The empirical results indicate classification as a public institution and percent of the financial support allocated to female athletes at an institution have a negative and statistically significant impact on six-year graduation rates of basketball programs. The two negative and statistically significant variables provide some interesting possible interpretations. First, the six-year basketball graduation rate for public institutions is over 15% lower than the comparable graduation rate at private institutions. The selectivity and higher admission standards at private institutions are likely contributors to the differential. Second, increasing the percentage of financial support for women's athletics appears to have an adverse impact on six-year graduation rates. The extra financial support toward women's teams is likely to have a positive impact on

the female graduation rate but the result implies the effort diminishes the male basketball graduation rate.

Profitability of the basketball program and average head coaches pay are both negative but not statistically significant determinants of college basketball six-year graduation rates. The lack of significance in the program profitability variable provides anecdotal evidence supporting the hypothesis that college athletics for high profile teams and athletes may have a propensity to focus more on placement at the professional sports level than earning a college degree.

One avenue for future research is to see if the empirical results are consistent across other sports with both men's and women's teams. Track and field teams could serve as an ideal sport for a comparison to basketball. A second approach for future research is to investigate the determinants of six-year graduation rates of athletic programs taken together instead of focusing on specific sports. Capturing college football as part of an aggregate effort is an important financial driver for many athletic programs.

REFERENCES

Astin, A. & L. Oseguera (2002). *Degree Attainment Rates at American Colleges and Universities*. Los Angeles, UCLA Graduate School of Higher Education Research.

Baade, R., R. Baumann & V. Matheson (2008). Assessing the economic impact of college football games on local economies. *Journal of Sports Economics*, *9*(6), 628-643.

Brown, C. & J. Medoff (1989). The employer size-wage effect. *Journal of Political Economy*, *97*(5), 1027-1059.

Brown, K. & L. Jepsen (2009). The impact of team revenues on MLB salaries. *Journal of Sports Economics*, *10*(2), 192-203.

Fox, J. T. (2009). Firm size wage gaps, job responsibility, and hierarchical matching. *Journal of Labor Economics*, *27*(1), 83-126.

Holmes, J., J. Meditz & P. Sommers (2008). Athletics and alumni giving: Evidence from a highly selective liberal arts college. *Journal of Sports Economics*, *9*(5), 538-552.

Humphreys, B. R. (2000). Equal pay on the hardwood: The earnings gap between male and female NCAA division I basketball coaches. *Journal of Sports Economics*, *1*(3), 299-307.

Idson, T. & L. Kahane (2000). Team effects on compensation: An application to salary determination in the national hockey league. *Economic Inquiry*, *38*(2), 345-357.

Kahn, L. M. (2006). Race, performance, pay, and retention among national basketball association head coaches. *Journal of Sports Economics*, *7*(2), 119-149.

Leeds, M. & P. Von Allen (2002). *The Economics of Sports*. New York: Addison Wesley.

Lentz, B. & D. Laband (2009). The impact of intercollegiate athletics on employment in the restaurant and accommodations industries. *Journal of Sports Economics*, *10*(4), 351-368.

Lucas, R. E. (1978). On the size distribution of business firms. *Bell Journal of Economics*, *9*(2), 508-523.

Mangold, W., L. Bean, & D. Adams. The impact of intercollegiate athletics on graduation rates among major NCAA Division I universities: Implications for college persistence theory and practice. *The Journal of Higher Education*, *74*(5), 540-562.

McEvoy, C. (2005). The relationship between dramatic changes in team performance and undergraduate admissions applications. *SMART Journal*, *2*(1), 17-24.

Metzger, B. & J. Bean (1987). The estimation of a conceptual model of nontraditional undergraduate student attrition. *Research in Higher Education*, 27(*1*), 15-38.

Monks, J. (2003). Patterns of giving to one's alma mater among young graduates from selective institutions. *Economics of Education Review*, *22*(2), 121-130.

Oi, W. (1983). Heterogeneous firms and the organization of production. *Economic Inquiry*, *21*(2), 147-171.
Pascarella, E. & P. Terenzini (1991). *How College Affects Students: Findings and Insights from Twenty Years of Research*. San Franciso, Jossey-Bass Publishing.
Rees, D. & K. Schnepel (2009). College football games and crimes. *Journal of Sports Economics*, *10*(1), 68-87.
Rhoads, T. & S. Gerking (2000). Educational contributions, academic quality, and athletic success. *Contemporary Economic Policy*, *18*(2), 248-258.
Rishe, P. (2003). A reexamination of how athletic success impacts graduation rates: Comparing student-athletes to all other undergraduates. *American Journal of Economics and Sociology*, *62* (2), 407-421.
Scott, M., T. Bailey & G. Kienzl (2006). Relative Success? Determinants of College Graduation Rates in Public and Private Colleges in the U.S. *Research in Higher Education*, 47(*3*), 249-279.
Siegfried, J., A. Sanderson & P. McHenry (2007). The economic impact of colleges and universities. *Economics of Education Review*, *26*(5), 546-558.
Smith, R. (2008). Big-time college basketball and the advertising effect. *Journal of Sports Economics*, *9*(4), 387-406.
Terry, N., R. Pjesky, & R. Patterson (2011). Determinants of men's college basketball profit. *Research in Business and Economics Journal*, *4*(1), 4-11.
Terry, N., R. Pjesky, & G. Rider (2009). Determinants of pay for head coaches in collegiate athletics. *The Business Review, Cambridge*, *13*(2), 66-72.
Terry, N. & C. Ramirez (2005). Gender equity regulation and profitability in college athletics. *Academy of Educational Leadership Journal*, *9*(3), 147-158.
Tinto, V. (1993). *Leaving College: Rethinking the Causes and Cures of Student Attrition*. Chicago, University of Chicago Press.
Tucker, I. (2005). Big-time pigskin success: Is there an advertising effect? *Journal of Sports Economics*, *6*(2), 222-229.
Turner, S., L. Meserve & W. Bowen (2001). Winning and giving: Football results and alumni giving at selective private colleges and universities. *Social Science Quarterly*, *82*(4), 812-827.
U.S. Department of Education, National Center for Education Statistics (2012). *The Condition of Education 2011* (NCES 2012-045), Indicator 45.

ASSET ALLOCATION BASED ON ACCUMULATED WEALTH AND FUTURE CONTRIBUTIONS

William J. Trainor Jr., East Tennessee State University

ABSTRACT

The median accumulated investment balance for investors with 10 to 15 years to retirement falls drastically short of what is needed with some studies suggesting more than half the population in this age group have virtually zero savings. Individuals who find themselves in this predicament and intend to make near certain future contributions should consider the present value of these future contributions as a risk-less income stream into their retirement account. With this in mind, early contributions should generally be directed towards 100% equity or similar risk-return asset classes. Using a simplified 50/50 stock-bond example, adjusting contributions to account for this unrealized stream of "risk-free" cash into the retirement account will increase expected terminal wealth after 15 years by approximately 10% with minimal increases in end of horizon risk, although within-horizon risk is magnified. For those with significant balances, consideration of future contributions is not as critical.

INTRODUCTION

The typical asset allocation model almost exclusively focuses on the risk/return relationship for assets already realized and invested. There is the classic age in bonds or 100 minus age model to determine the percentage in equities. A variety of target date or life cycle funds follow this type of concept. Other models are based solely on an investor's risk aversion and will often delineate portfolios as conservative, moderate, or aggressive. Regardless of the asset allocation model used, two major considerations are invariably overlooked: current wealth and expected future contributions. Friend and Blume (1975) first pointed this out and even stated, "virtually all empirical applications of portfolio theory have ignored human wealth in spite of its obvious importance to the demand for risky assets." This issue remains in the financial planning area to this day.

To explain further, consider two types of investors with 15 years until retirement. Each earns $50,000 a year and both plan to make $10,000 contributions each year. However, the first investor has zero invested wealth while the second has already accumulated $300,000. A typical age rule might suggest a 50/50 mix. However, this is biased downward for both investors if future contributions are not considered.

Although the first investor has zero accumulated wealth, there is $150,000 in "riskless" future contributions. This is riskless only in the sense that it is assumed the investor will not lose

his or her income stream, face unexpected expenses, etc. that would derail future planned contributions. Thus, for the first few years, this investor may want to consider 100% in equities until wealth at risk relative to future contributions has increased. Large losses at this point in the accumulation phase, despite the late start and limited time horizon, will be mitigated by future contributions. In addition, future social security payments which can also be considered a risk-free annuity will be a much larger proportion of retirement income, further increasing the actual percentage of wealth in relatively risk-free low yielding assets.

For the second investor, wealth at risk is much greater as future contributions are only 50% of accumulated value. However, the investor still has 33% in "riskless" future contributions, $150,000/$450,000. At this point, a true 50/50 mix would mean the investor should have $225,000 in equities. Depending on asset returns and how the suggested asset allocation adjusts through time, this investor's initial contributions will be directed towards both bonds and stocks although not at the implied 50/50 ratio.

Thus, financial planners and investors need to focus not just on the risk/return relationship for assets in the retirement account, but also need to account for those assets that have not yet been earned, but will be directed towards retirement. This study shows the risk/return characteristics of the classic investing approach versus considering the inclusion of future contributions. Findings suggest with little difference in terminal risk, expected terminal wealth could be increased by approximately 10% for investors with no accumulated balances. For those with significant balances, the consideration of future contributions is not as critical.

TARGET INVESTMENT GROUP

Although this analysis can be effectively applied to any investor at any age, it is likely more relevant to investors that have greater certainty about future contributions. This would seem to be particularly apt for investors in the 50 to 65 age group category as their children are likely out of college, income is peaking, retirement savings have become a priority, and on average, there is less uncertainty about job security. These factors should lead to greater certainty about what can and will be contributed towards retirement.

Unfortunately, many investors even at this age have little savings. Recent news based on a variety of surveys (Employee Benefit Research Institute (EBRI), Fidelity, Federal Reserve) place median retirement account values for those between 45-65 anywhere from $65,000 to $120,000, (American Association of Retired Persons (AARP), 2013; Average Retirement Savings Guide, 2013; Greenhouse, 2013). The results of many of these surveys are likely biased upwards just based on the clientele surveyed. The Schwartz Center for Economic Policy Analysis (SCEPA) using 2010 Census data estimated that 75% of those in the 50-64 age (43 of 58 million) have a paltry median retirement savings of $6,500, see Table 1.

Table 1 Estimated retirement balances based on analysis and surveys from SCEPA, EBRI, Fidelity, and the Federal Reserve		
SCEPA, Age 45-65	Retirement Balances	
Income	Median	Mean
Bottom 25%: $0 to 10,800	$0	$16,034
25-50%: $10,801 to $27,468	$0	$21,606
50-75%: $27,469 to $52,200	$6,500	$41,544
75-100%: $52,201+	$52,000	$105,012
EBRI, Age 50-54	NA	111,900
Fidelity, 55-64	NA	$65,000
Federal Reserve, 55-64	NA	$120,000

This changes the dynamic of contribution value to accumulated wealth. It is commonly suggested that at the age of 55, one should have saved approximately 5x their salary, (Fidelity; Greenhouse, 2013). With an income of $50,000, that would mean an investor should have $250,000 at this point. For investors with little or no accumulated wealth, focusing on the target date risk/return tradeoff instead of the instantaneous risk/return tradeoff will improve the expected ending outcome while adapting to risk preferences of the individual as quickly as possible. In addition, this method avoids the use of margin and leverage which few beginning investors are likely to employ, nor are even able to if using standard work related 401k accounts.

INVESTMENT PLANNING AND RISK AVERSION

Several studies have suggested investors have constant relative risk aversion led by Friend and Blume, (1975). Thus, regardless of wealth, the percentage held in the risky asset would remain the same. Most of the studies that come to this conclusion are based on cross sectional data and do not give any indication how relative risk aversion may change for an individual with changes of wealth. Guiso and Paiella (2008) conclude there is decreasing absolute risk aversion as wealth increases while Chiappori and Paiella (2011) conclude there could be decreasing relative risk aversion depending on the underlying assumptions. If this is the case, and allowing the additional assumption that wealth for retirement includes future contributions, then the amount of equity exposure should be higher than usually considered prudent. For those with no accumulated wealth, this increase can be dramatically higher.

Using Friend and Blume's (1975) derivations, it is easy to show how the amount in the risky asset is a function of wealth. They show the following:

$$E(R_m - r_f)/\sigma^2_m = C^*[R/W + \beta_{hm^*} H/W] \quad (1)$$

where $E(R_m - r_f)/\sigma^2_m$ is the expected risk premium per unit of risk for the market portfolio, R is the liquid wealth amount in the risky asset, β_{hm} is the ratio between the market covariance and an investor's human capital divided by the variance of the market, H is the value of the investor's human wealth, W is the sum of all wealth, and C is the risk aversion parameter. Their study involved estimating C, while R/W is of concern here. Setting $\beta_{hm} = 0$ which assumes future contributions are not correlated to market returns, C = 2 which is the general average estimate of risk aversion, $E(R_m - r_f) = 0.05$ which is the approximate historical equity risky premium, and $\sigma^2_m = 0.04$ which corresponds to a 20% market standard deviation, equation (1) simplifies to:

$$R/W = 0.55. \qquad (2)$$

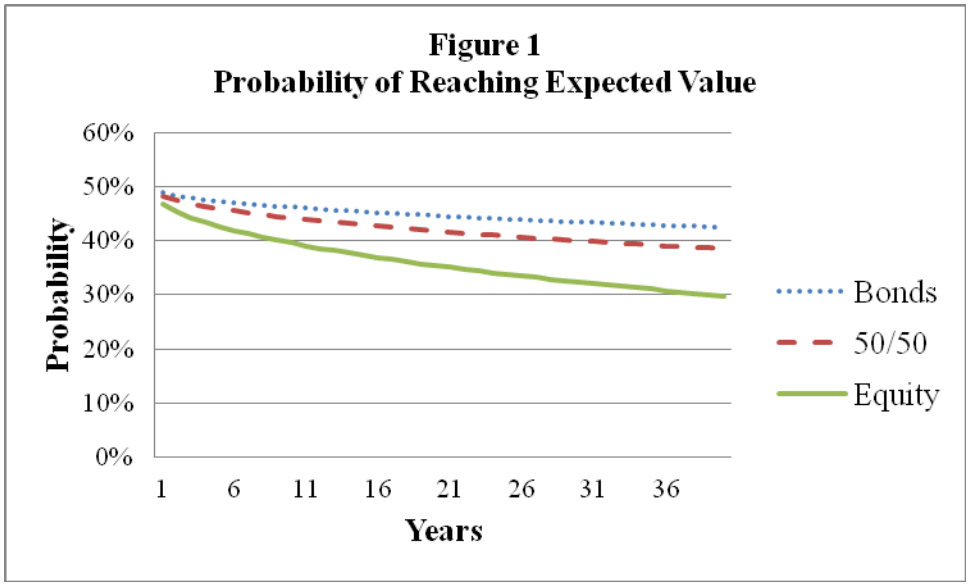
This suggests 55% of total wealth should be invested in the risky asset. Including human capital as part of wealth, the actual amount of liquid wealth invested in the risky asset increases. As an example, assume human capital is ignored and the investor has $50 of liquid capital. 55% of this is $27.50 which would be invested in the risky asset. If another $50 in human capital or future contributions is expected, 55% of $100 is $55, suggesting 100%+ of liquid wealth should be in the risky asset. If C is indeed related to wealth suggesting decreasing relative risk aversion, an even greater percentage in the risky asset would be calculated. Including human wealth defined here as the present value of future contributions suggests a greater percent of realized or liquid wealth should be in the risky asset. Adding the present value of social security payments to wealth would further increase the equity percentage of liquid wealth.

DATA AND METHODOLOGY

The Center of Research and Security Prices (CRSP) S&P 500 value weighted index is used as a proxy for monthly equity returns. 10-year Treasury returns are used as a proxy for bond returns, although prior to May 1941, 90-day T-bill returns are used as 10-year data is not available before 1941. Data covers the Jan. 1926 to the Dec. 2012 time frame. Since this study looks at a 15-year horizon, data is limited even using overlapping monthly time periods. In an effort to project what may occur, while still maintaining the correlation between bond and equity returns, along with any intertemporal correlation among stock and especially bond returns, bootstrapping is employed. To create 15 years of monthly returns, data is re-sampled 6 months at a time with replacement from 1044 months of historical data that is available. This still leaves 1039 overlapping 6 month periods to sample from. 10,000 simulations are employed for each run resulting in trivial differences between separate 10,000 runs.

Both means and medians of terminal wealth are reported. This is particularly relevant when dealing with compounded returns as they are lognormally instead of normally distributed. This creates the situation when the probability of reaching the mean is much less than 50%, (Booth, 2004). Figure 1 shows these probabilities based on the data. For an all equity portfolio,

the probability of actually reaching the mean in 15 years is less than 40%. Thus, medians may be the more relevant statistic to consider.



This also makes the interpretation of the standard deviation problematic and thus, standard deviations are not reported. To give a more accurate picture of the risk, 90% confidence intervals are given to indicate the range of values. In addition, within horizon risk is also measured. Both the median worse loss during the 15 years is shown, along with 90% confidence intervals for this loss. This becomes relevant for any investor who is wary of large losses and may exit the market during non-fortuitous times. Mutual fund flows suggests this behavior is common as investors tend to exit out of funds after they drop in value.

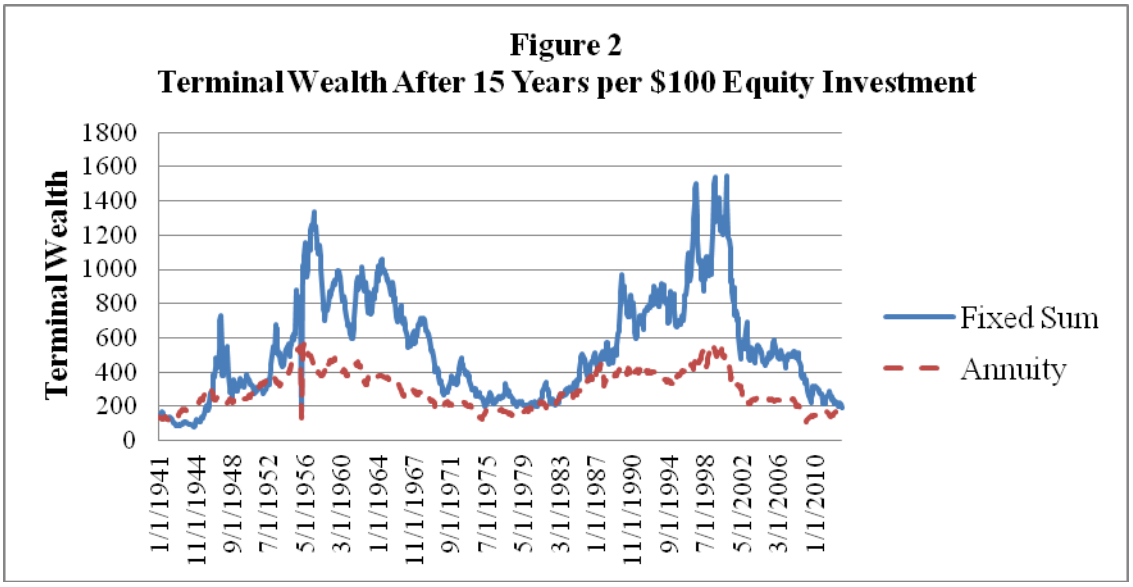
EMPIRICAL RESULTS

Standard analysis usually compares lump sum investing to dollar cost averaging. The literature is fairly extensive in this area, with most all studies agreeing lump sum investing results in the highest expected value, although depending on the particular return path, dollar cost averaging can result in higher returns, (Constanides, 1979; Knight and Mandell, 1993; Williams and Bacon, 1993; Rozeff, 1994; Israelson 1999; Abeysefera and Rosenbloom 2000; Leggio and Lien, 2003, Milevsky and Posner 2003 to name a few). Studies tend to show that the advantage of dollar cost averaging is less risk, especially when it comes to within horizon risk (Dubil 2004; Trainor, 2005)

A. Historical Results

Figure 2 shows the historical results for a 15-year horizon in comparing lump sum to dollar cost averaging with one caveat. That is, for the annuity, the present value of $100 is spread out over the 15-year horizon using a 2% discount rate. The annuity itself increases once a year by 2%. This is to show the difference between someone who has $100 to invest right now versus someone who has just started to contribute. Thus, the results here may differ quantitatively from most studies since it is usually assumed that the amount not invested in equities is invested in some risk-free asset, and the time to dollar cost average into the market is generally assumed to occur over a much shorter time-period, usually 2 to 5 years.

Figure 2
Terminal Wealth After 15 Years per $100 Equity Investment



As expected, having $100 and immediately investing the whole sum in equities results in much higher values of terminal wealth. Figure 2 also demonstrates how critical the start date can be showing the beginning of the early 1940's and 80's being very profitable. Annuity values are much less variable as might be expected when investing over a 15-year period. At the very least, if one does indeed have a large fixed sum, dollar cost averaging over 15 years has rarely paid off relative to investing in stocks immediately.

Table 2 demonstrates how skewed these values can be with the medians significantly less than the mean. Within horizon risk is also significant for fixed sum investors with an average loss of 13% sometime during the 15 year period and a lower confidence interval limit of being down 69%. This means there is still a 5% chance of being down more than -69% at sometime during the investment period. As expected, bonds have the least risk reducing the average within horizon loss to -3%, with a lower confidence interval level of 13%.

Table 2
Overlapping monthly 15 year periods from Jan. 1926 to Dec. 2012. Terminal Values per $100 invested.

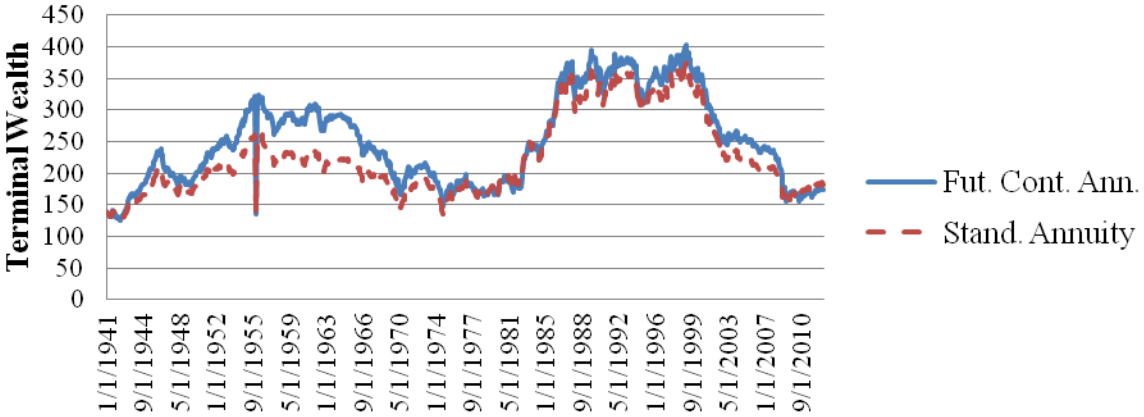
	Median	Mean	Lower limit	Upper Limit	Median Within Horizon Loss	Within Horizon lower CI
Fix Sum stocks	$482	$542	$136	$1,112	-13%	-69%
Fix Sum bonds	$180	$237	$110	$503	-3%	-13%
Fix Sum 50/50	$307	$345	$154	$681	-7%	-39%
Annuity 100% stocks	$272	$296	$141	$491	-1%	-16%
Annuity 100% bonds	$162	$177	$121	$280	0%	-1%
Annuity 50/50	$212	$227	$154	$350	0%	-8%
Future Contribution Annuity 50/50	$244	$253	$164	$374	-1%	-13%

Dollar cost averaging into any asset category dramatically reduces both the median and means along with increasing the lower limits. Within horizon risk is significantly reduced with median loss values down only 1%. As an example, there is only a 5% chance of being down more than -16% in stocks relative to the total amount planned on being invested.

The primary focus of this study is the comparison of the investor who desires a 50/50 mix and makes steady contributions over the 15-year period. The standard asset allocation approach which involves investing contributions into a 50/50 fund has a historical median of $212 and a lower confidence level limit of $154. Considering future contributions, this investor should consider investing 100% in equities for the first few years until a 50% balance is achieved between equities relative to bonds plus future contributions. This results in a mean of $244 and based on the lower limit, actually shows less risk, $164 versus $154. The within horizon loss is slightly higher than the standard annuity as there is a 5% chances of being down -13% or more compared to -8% with the standard annuity as shown in the last column.

Figure 3 shows the different outcomes based on the start date. Overall, using future contributions to determine an asset allocation actually reduces terminal risk while significantly increasing expected wealth along with the opportunity for larger outcomes.

Figure 3
Terminal Wealth After 15 Years per $100 Equity/Bond Investment



B. Monte Carlo Simulation

Although the historical results are informative, they are based on a single return path. To make sure the results are more robust, bootstrapping is employed. Table 3 shows the results.

Table 3
Comparison of Fixed Sums to Steady Contributions per $100 Total Investment.

	Median	Mean	Lower CL	Upper CL	Median Within Horizon Loss	Within Horizon lower CI
Fix Sum stocks	$401	$522	$108	$1,316	-14%	-50%
Fix Sum bonds	$203	$212	$134	$321	-3%	-14%
Fix Sum 50/50	$298	$321	$151	$562	-5%	-23%
Annuity 100% stocks	$256	$290	$113	$579	-5%	-22%
Annuity 100% bonds	$164	$168	$130	$216	-1%	-1%
Annuity 50/50	$211	$218	$137	$323	-1%	-5%
Future Contribution Annuity 50/50	$231	$243	$133	$393	-4%	-16%

Simulation results are similar to the historical results although there are some significant differences, especially with stock returns. Focusing on the 50/50 mix and considering future contributions and allocating assets accordingly again leads to an approximately 10% increase in the mean and median ending wealth values. Unlike the historical results, the lower limit is indeed lower when considering future contributions, but not dramatically. The within horizon risk is much larger with a 5% probability of being down -16% or more as opposed to only -5%

with the standard annuity. Thus, although there is a 10% increase in terminal wealth for very similar end of horizon risk, within horizon risk needs to be considered. This risk is especially relevant for investors that may either stop contributing or change the asset allocation significantly if large losses occur during the time horizon.

Finally, Table 4 displays results assuming the investor starts with 3x the present value of the future contributions. In this case, the difference in terminal wealth is less than 4% and given the higher risk, does not imply that the consideration of future contributions to determine asset allocation is critically important. Thus, if investors do have large investment balances relative to future contributions, they should indeed primarily focus on the risk/return relationship of their actual retirement account balance.

Table 4
Comparison of Fixed Sums to Steady Contributions with $300 initial, $100 additional for Annuity.

	Median	Mean	St. Dev.	Lower CL	Upper CL	Median Within Horizon Loss	Within Horizon lower CI
Annuity 100% stocks	$1,521	$1,902	$1,527	$458	$4,633	-12%	-42%
Annuity 100% bonds	$766	$799	$191	$541	$1,149	-2%	-7%
Annuity 50/50	$1,148	$1,221	$470	$613	$2,091	-4%	-18%
Future Contribution Annuity 50/50	$1,188	$1,274	$525	$599	$2,251	-6%	-24%

CONCLUSION

Investors do not appear to be saving enough to avoid a drastic drop in their standard of living at retirement. In fact, SCEPA research based on 2010 census data show that the median investment savings for 75% of people aged 50-64 is $6,500. Less dire survey results still suggest this value is no more than $100,000. Investors in this situation face difficult choices with only 10-15 years until retirement.

Typical average risk aversion parameters at this stage in life may suggest anywhere from a 40-60% exposure to stocks. However, with little or no accumulated savings at this point, placing future contributions into a 50/50 stock/bond portfolio may not be the optimal choice. At this stage, it would seemingly be expected that investor's future retirement contributions have become a priority. Assuming these contributions are relatively certain results in an interesting dichotomy between those who have significant accumulated savings and those that do not.

Treating future contributions as if they were sitting in a risk-free asset and earning a rate of return equal to any future increase in the value of the contribution, means that investors just starting to save have most all of their projected future retirement balance locked up in future contributions. Thus, this type of investor should seriously consider placing initial contributions in 100% equity or a similar asset class. Eventually, as the accumulated account balance equals

future contributions, current contributions can then be directed toward a more varied mix. For those with large account balances relative to future contributions, consideration of this issue is not as critical.

Both historical and Monte Carlo simulation suggests that this type of investment philosophy will result in a 10% increase in the expected account balance at retirement with little increase in terminal wealth, although within horizon risk is magnified. This drawback needs to be seriously considered. Individuals that do not have a history of investing may be more risk-averse and a 100% equity exposure to begin with, even though future contributions will minimize any early losses, may result in the investor leaving the market before more optimal results can be achieved.

REFERENCES

AARP, "401(k) Balances Hit Record High - $119,900 for Ages 50-54, posted 2/15, 2013, http://blog.aarp.org/2013/02/15/401k-balances-climb-to-record-high/.

Abeysekera, S. P., and E. S. Rosenbloom. (2000). A simulation model between lump sum and dollar-cost averaging. *Journal of Financial Planning*, 13, 86-92.

Arrow, K.J. (1965). The theory of risk aversion. Aspects of the theory of risk bearing, by Yrjo Jahnssonin Saatio, Helsinki. Reprinted in: Essays in the Theory of Risk Bearing, Markham Publ. Co., Chicago, 1971, 90–109.

Average Retirement Savings Guide. Retrieved May 2013 from http://www.averageretirementsavingsguide.com/.

Booth, L. (2004). Formulating retirement targets and the impact of time horizon on asset allocation, *Financial Services Review*, 13, 1-17.

Chiappori, Pierre-Andre, and M. Paiella. (2011). Risk Aversion is Constant: Evidence from Panel Data. *Journal of the European Economic Association*, 9(6), 1021{105

Constantinides, G. M. (1979). A note on the suboptimality of dollar-cost averaging as an investment policy. *Journal of Financial and Quantitative Analysis*, 14, 443-450.

Dubil, R. (2004). The risk and return of investment averaging: An option-theoretic approach. *Financial Services Review*, 13, 267-283.

Milevsky, M. A., and S. E. Posner. (2003). A continuous-time reexamination of dollar-cost averaging. *International Journal of Theoretical and Applied Finance*, 6, 173-194.

Fidelity, How Much Do You Need to Retire. 2/27/2013, Retrieved May, 2013, from https://www.fidelity.com/viewpoints/personal-finance/8X-retirement-savings.

Friend, I. and M. Blume. (1975). The demand for risky assets. *American Economic Review*, 65, 900-922.

Greenhouse, S. (2013). Pushing back retirement, and not always for money, NY Time, March 12, 2013, Retrieved May (2013) from http://www.nytimes.com/2013/03/13/business/retirementspecial/pushing-back-retirement-and-not-always-for-money.html?pagewanted=all&_r=0.

Guiso, L. and M. Paiella. (2008). Risk aversion, wealth, and background risk. *Journal of the European Economic Association*, 6: 1109–1150.

Israelsen, C. L. (1999). Lump sums take their lumps. *Financial Planning*, Jan. 51-60.

Knight, J. R., and L. Mandell. (1993). Nobody gains from dollar cost averaging: Analytical, numerical and empirical results. *Financial Services Review*, 2, 51-61.

Leggio, K. B. and D. Lien. (2003). An empirical examination of the effectiveness of dollar-cost averaging using downside risk performance measures, *Journal of Economics and Finance*, 27, 211-223.

Pratt, J. W., (1964). Risk aversion in the small and in the large. *Econometrica* 32, January–April, 122–136.

Rozeff, M. S. (1994). Lump-sum investing versus dollar-averaging. *The Journal of Portfolio Management*, 4, 45-50.

Save, spend, splurge, how much should you have saved for retirement by Age? March, 26th, 2013, retrieved May, 2013 from http://www.savespendsplurge.com/2013/03/26/how-much-should-you-have-saved-for-retirement-so-far-by-age/.

Schwartz Center of Economic Policy Analysis. Retirement account balance by income: Even the highest earners don't have enough. Retrieved May, 2013 from http://www.economicpolicyresearch.org/guaranteeing-retirement-income/528-retirement-account-balances-by-income-even-the-highest-earners-dont-have-enough.html.

Trainor, W. J. (2005). Within-horizon exposure to loss for dollar cost averaging and lump sum investing," *Financial Services Review*, Vol. 14, Is. 4, pp. 319-330.

Williams, R. E. and P. W. Bacon. (2004). Lump sum beats dollar cost averaging. J*ournal of Financial Planning*, 17, 92-95.