ACADEMY of ECONOMICS and FINANCE

PAPERS AND PROCEEDINGS

39th
ANNUAL MEETING

CHARLESTON, SOUTH CAROLINA
FEBRUARY 8 – 11, 2012

editors
DR. ADAM JONES
DR. CLAY MOFFETT
UNC WILMINGTON

AEF Papers and Proceedings, Volume 35

Academy of **Economics**
and **Finance**

Papers and Proceedings

Volume 36, 2012

Thirty-ninth Annual Meeting
Charleston, South Carolina
February 8-11, 2012

Program Arranged by Dr. Rob Burrus
University of North Carolina Wilmington

Editors

Dr. Adam T. Jones and Dr. Clay M. Moffett
University of North Carolina Wilmington

2012-2013 AEF Leadership

President Robert Burrus -- University of North Carolina Wilmington

President Elect Jocelyn Evans -- College of Charleston

First Vice-President Doug Berg -- Sam Houston State University

Second Vice-President Ed Graham -- University of North Carolina Wilmington

Secretary Albert DePrince, Jr. -- Middle Tennessee State University

Treasurer Richard M. Vogel -- State University of New York at Farmingdale

Directors

Steve Johnson -- Sam Houston State University (2013)

Michael McLain -- Hampton University (2013)

Mary Kassis -- University of West Georgia (2014)

Shahdad Naghshpour -- University of Southern Mississippi (2014)

Stephen Conroy -- University of San Diego (2015)

Speros Margetis -- University of Tampa (2015)

Olivier Maisondieu Laforge -- University of Nebraska Omaha (2016)

Clay M. Moffett -- University of North Carolina Wilmington (2016)

AEF Papers and Proceedings, Volume 35

Recent Presidents

Steven Jones -- Samford University

Robert Stretcher -- Sam Houston State University

Richard McGrath -- Armstrong Atlantic State University

Bill Sackley -- University of North Carolina Wilmington

David Boldt -- University of West Georgia

Albert DePrince, Jr. -- Middle Tennessee State University

Graham Mitenko -- University of Nebraska - Omaha

Bichaka Fayissa -- Middle Tennessee State University

David Lange -- Auburn University Montgomery

Richard Cebula -- Armstrong Atlantic University

David Sollars -- Auburn University Montgomery

William W. Hall, Jr. -- University of North Carolina at Wilmington

James Lindley -- University of Southern Mississippi

Editors, AEF Papers and Proceedings

Adam T. Jones and Clay M. Moffett -- University of North Carolina Wilmington

Editor, Journal of Economics and Finance

James Payne -- University of South Florida Polytechnic

Editor, Journal of Economics and Finance Education

Richard Cebula -- Armstrong Atlantic University

Editors, AEF Journal

Robert T. Burrus, Jr. and J. Edward Graham Jr. -- University of North Carolina Wilmington

Webmaster

Doug Berg -- Sam Houston State University

Academy of Economics and Finance

Volume 36 **Papers and Proceedings** **2012**

Table of Contents

AEF Papers and Proceedings, Volume 35

AEF Papers and Proceedings, Volume 35

Occam In October

Ronald E. Crowe, Our Lady of the Lake University
George Ignatin, University of Alabama - Birmingham

Abstract

Major League Baseball plays the World Series every October. We are trying to determine the best, simple (Occam's razor) measure of the number of runs each team will score. We examine this measure's relationship to salaries and World Series Champions. We test hypotheses on recent World Series data that Slugging percentage times On-base percentage, SLOB, is the best simple measure of producing runs, that teams pay for SLOB, and that SLOB helps predict World Series Champions.

Introduction

Pluralitas non est ponenda sine necessitate – William of Occam
Good pitching will beat good hitting any time, and vice versa. Bob Veale, 1966

A scene from the movie Moneyball (2011) depicts a dilemma between subjective observation of baseball skills and the objective nature of statistical analysis. This was the story behind Michael Lewis's book of the same name (2003). Can a statistic predict results? Do teams pay for the statistic? Does the statistic predict a World Series win?

The month of October is the culmination of the entertainment and athletic achievement of thirty-two professional teams of Major League Baseball; the World Series of Major League Baseball. What will help the team achieve four victories in seven games? We know that the team wins with the most runs scored in a game; scoring by moving a batter-runner around the four bases. Baseball is unique in that control of offense is in the arm of the pitcher; the batter is dependent initially on the pitches.

We know that teams pay salaries and players perform. *USA Today* (2011) data reveal that the median payroll for the 2011 season was $86,763,646. Each team has a roster of 25 players. Non-pitchers make up slightly less than half of the team with a median salary just over one million dollars per season. They account for 57%, on average, of a team's salaries. We assume that the remaining salary is for pitching. Except where premium ticket pricing takes place, the data to calculate marginal revenue product of players are proprietary. We focus on the marginal product of batting and choose four statistics to predict runs.

Measurement of Marginal Productivity

Economists usually assume that workers are homogenous, and that when one employs more of these variable inputs with a fixed amount of capital, their Marginal Physical Product (MPP) will diminish. This abstraction obviously is not a realistic assumption but provides a useful starting point for the production of manufactured goods. It is even less realistic and less useful for analyzing professional sports. In the case of baseball, we have two productive activities: offense and defense. We divide production further; between pitching and fielding on defense, and batting and running on offense, respectively. Regardless of the intensity of the talent and skill necessary to produce, in the final analysis, winning is the goal with runs therein as the means. Baseball is unique among team sports in that offense and defense are completely independent. In addition, there is a remarkable quantity of data available to study.

The purpose of our paper is to examine three items: the conceptual and empirical issues of offense, namely hitting or coincidently the opposing pitching; the payment of the resource; and predictions of the World Series winner. We look at previous descriptions and conceptual issues. Then, we empirically test the four simplest measures of pitching and batting against runs per game, salary measures, and predicting the champion.

Measurement Issues

Batting Average

The oldest and simplest metric for hitting is Batting Average (BA) or the number of hits divided by the number of official at bats (AB). 1. See Equation 1.

$$BA = \frac{H}{AB} \quad (1)$$

$$WHIP = \frac{H + BB}{IP} \quad (2)$$

$$\text{OBP} = \frac{\text{H} + \text{BB} + \text{HBP}}{\text{AB} + \text{BB} + \text{HBP} + \text{SH} + \text{SF}} \quad (3)$$

$$SLG = \frac{1 \times 1B + 2 \times 2B + 3 \times 3B + 4 \times HR}{AB} \quad (4)$$

Theoretically, the measurement leaves out the quality of the hit and getting on base by walk (BB), catcher obstruction, catcher error (not catching a called third strike), a pitcher's balk (illegal pitch with a runner on base), or being hit by a pitch (HBP). BA also does NOT count as an official AB for two special statistics: Sacrifice Flies (SF) and Sacrifice Hits (SH), i.e., bunts that advance a runner when the batter is, or should have been, put out. Extra base hits put hitters in better position to score than do singles and, probably more important, move runners on base farther than do singles, BB or HBP. However, a walk does productive things for the team. First, it avoids an out, secondly, it puts a runner in position to score, and lastly, it puts up another player to bat.

What makes a measure of quality? If we use a measure based on at bats, e.g. runs per at bat, this presents an issue of advantage. When a batter advances from the batter's box to first base, it allows another batter to the plate. In other words, any "at bat" ratio tends to overstate the poorer team and understate the better team on offense.

Walks plus Hits per Inning Pitched, On Base Average, and Slugging Percentage

Walks plus Hits per Inning Pitched, WHIP, is a pitching statistic. It captures more than the batting average but less that On Base Percentage. See equations 2 and 3. WHIP is a measure of pitchers' performance and ties closely to a game's action. If a home team is ahead in the ninth inning after the visiting teams half (3 outs recorded), the home team wins and has one less inning of pitching. When comparing two teams, there should not be much difference between team A's WHIP and team B's On Base Average.

On Base Average (OBA) or On Base Percentage (OBP) measure all means to get on base divided by all appearance at the plate. See equation 3. Thus, OBP eliminates several of the conceptual problems of BA; but it does not solve the "quality" of hits problem, most of which is solved by Slugging Percentage, SLG. See equation 4. There are two problems with SLG: first, it does not take into account BB and HBP, and second, it overstates the value of extra-base hits. Although HR's are more productive than singles, they are not four times as productive. Note that a player who hits one HR every four AB's has a 1.000 SLG as would a player who had four singles in four AB. Obviously, four singles in any one inning create more runs or scoring opportunities than a HR, but a HR with the bases loaded creates more runs than four lone singles in different innings. In other words, in games with no extra-base hits, slugging is the same as batting average. With few extra base hits and the small size of the observations in a world series, slugging and its complements (OPS and SLOB) may not have the best fit.

Because OBP and SLG have complementary strengths and weaknesses and together solve most of the conceptual problems with BA, it has dawned on several writers that some combination of the two would be better than either one by itself. As with most insights into baseball statistics, the seminal thinker in this area has been Bill James, who in the 1970's and 1980's wrote annually the *Baseball Abstract*. He argued that hitting had two important elements, which he imaginatively called "A" and "B". The A part was essentially OBP. The B part was essentially SLG.

OPS and SLO

Over the years, Bill James has produced a dizzying array of methods under the rubric "Runs Created." Based on the spirit of James' insight, Ignatin and Barra coined the term "SLOB," defined simply as SLG*OBP, in numerous articles in *The Village Voice*, *Inside Sports*, and *The Wall Street Journal*, dating from the mid 1980's. They argued that SLOB was the best measure of MPP. At about the same time, Tom Boswell, the excellent sportswriter for the *Washington Post*, developed a measure that he named "Total Average" which, as its name implies, incorporates almost every conceivable component of hitting in a formula that ADDS all of those components. Boswell does several things very well: his denominator is Outs, which is the right measure; he subtracts Caught Stealing (CS) from his numerator and adds CS to his denominator. For reasons of simplicity, we have grouped him with Palmer and Thorn, the editors of *Total Baseball*. The subtitle is *The Official Encyclopedia of Major League Baseball*, but might well be "The Bible of Baseball Stats." Palmer and Thorn coined the term

"OPS," defined simply as OBP plus SLG, which they say is the best measure of hitting production. Major League Baseball broadcasts of games by ESPN and FOX networks include the listing of OBP.

The Impetus for Occam's Razor

Which is the better measure: SLOB or OPS? To answer that question, we have run regressions for MLB for the last ten World Series. The first problem we had was to define the proper dependent variable for Productivity. There are several: Runs (R) per game, R/AB, R/Outs, and R/Inning. R/AB seems like a good measure but has an implicit difficulty: the higher the On Base Average (OBA or OBP), the more AB. Thus, the better teams will have more AB/G than the worse teams. R/Outs and R/Inning are the best conceptual measures, but the World Series data are a relative small sample to obtain testable results.

Because of these problems, we use R as the appropriate dependent variable. However, there are two problems with this measure. Teams may play a different number of innings per game because not every team comes to bat for nine full innings. Some games go into "extra innings," while a few others are shortened by "environmental" problems. Finally, home teams (HT) do not bat in the ninth inning of games in which they already have a lead, and do not get their full three outs whenever they score the winning run in the ninth or later inning. Thus, by using R, we are assuming implicitly that these problems are not severe enough to affect the data. Before we discuss the results, we look at the literature.

Previous Attempts: Descriptions and Conceptual Issue

Economic Literature and the Complex

We start with a review of some of the descriptions of MRP in the literature. Davenport (1969) contends that marginal product of a baseball player is difficult to measure because the sport is team based. That is, a great player statistically and popularly is circuitously affected by the performance of others on the team. Davenport used the additional ticket sales generated when Sandy Koufax pitched for the Los Angeles Dodgers as a direct measure of the marginal revenue product. He infers the direct value of Koufax appearances, but not the actual statistics. It would be remiss not to suggest that the attendance is dependent on the successful performance statistics of Koufax, the Dodger offense, the Dodger defense, or all of these.

However, Davenport does estimate the indirect value of Koufax with at least one performance statistic by means of attendance differences with the last year Koufax played (1966) and the next year after he retired (1967). The statistic is his win-loss record compared to the average Dodger pitcher in the absence of Koufax. One tangent on player statistical comparisons: the next best in quality to Koufax would perhaps have been Don Drysdale, the next best pitcher.

Scully (1974) with offensive and defensive statistics, first estimates the percentage wins as a function of team slugging, team strike out-walk ratio, indicators for the National League, 20 games or more behind, and pennant race participants. Secondly, he specifies revenue as a function of percentage wins, the community population, value of tickets and broadcasts, percentage of African-American ball-players, and indicators on the National League and stadium age. There is an estimate of the marginal revenue product. It is the coefficients in the former equation for team slugging or team strike out-walk ratio, respectively for hitters or pitchers, times the value of the coefficient for percentage wins in the latter equation. One conceptual issue we have with the use of slugging is that it will miss the quality of getting on base as previously mentioned. As for empirical results, the first equation had an R^2 of 0.88 and the second, 0.75. Somers and Quinton (1982) take a variant of Scully's model with an R^2 of 0.89.

Some econometric models use the Cobb-Douglas production function. Hech (1981) states: "The obvious measure of a player's ability to hit frequently is his batting average, while homeruns provide a good measure of a player's ability to hit with power."

He used natural logarithmic transformations of variables to estimate wins. The empirical results showed that "hitting on average contributes almost 6 times as much as pitching [as measured by strike-out to walks] (p. 21)." Further, homeruns were two times greater than stolen bases towards wins. Hech's four varied models had R^2s that ranged from 0.773 to 0.790.

Woolway (1997), also with a Cobb-Douglass model, used slugging percentage, stolen bases, earned run average, stolen bases, and unearned run average. The coefficients, except for earned run average, were statistically different from zero. His results were consistent with Hech: batting - the slugging percentage - was more important than pitching.

We conclude that the economic models do well at predicting wins. The selection of the particular batting metrics all lack in terms of quality, e.g. batting average, homeruns and slugging.

Popular Baseball Research

Next, we looked at the popular baseball research area, the popular press and baseball enthusiasts. Bill James (1984) presented what he called the Pythagorean Theorem on the relationship between offensive runs scored (ORS), defensive runs allowed (DRA) and win percentage. Thus

$$\frac{Wins}{Game} = \frac{ORS^2}{(ORS^2 + DRA^2)} \qquad 5$$

James adjusts his model by making the exponent 1.83. A logarithmic transformation could estimate a coefficient for the exponent. James' model does take into account the offensive and defensive efforts but not specific marginal products of players.

In a different model, James's Run Created model takes into account hits, walks, hit by pitch, caught stealing, grounded into double play, intentional walks, sacrifice hits, sacrifices flies, total bases, and strikeouts. Conceptually, these could happen to a batter in a game, but the model lacks parsimony. Hence, its usefulness is doubtful.

Thorn and Palmer in *Total Baseball* (1994) specify the Total Player Rating (TPR). 4. The equation is

$$\text{Total Player Rating} = \text{ABR}+\text{FR}+\text{BSR} \qquad 6$$

where ABR is Adjusted Batting Runs, FR is Fielding Runs, and BSR is Base Stealing Runs.

It is probable that the coefficients are a result of estimation with linear regression. We acknowledge that equations ABR and FR have value. In our investigation, except for Ricky Henderson, a well-known "stealer," or the 1985 St. Louis Cardinals, also well known for stealing bases, rarely do stolen bases contribute to runs.

Boswell (1981) uses the Total Average from *Inside Sport Magazine* to evaluate major league players. The measure is

$$\text{Total Average} = \frac{(\text{singles} + 2\times\text{doubles} + 3\times\text{triples} + 4\times\text{homeruns} + \text{hit by pitch} + \text{base on balls} + \text{stolen bases} - \text{caught stealing})}{(\text{at bats} - \text{hits}) + \text{caught stealing} + \text{doubleplays})} \qquad 7$$

The numerator represents offense and the denominator outs. The Total Average is a measure of comparison of one player to another. It will tend to give more weight to the power hitter.

Ignatin and Crowe (2005) test the hitting metrics versus runs per game as the measure of the marginal product of major league batting by season, cross-sectional with teams. Lahman (2004) compiled these data. They find that OPS and SLB produce observably higher R^2s and F-statistics than BA, OBP, and SLG for teams by season from 1874 to 2004. There is a statistical difference between the R^2s and the F-statistics (Ignatin & Crowe, 2005).

Economics: SLOB, Salary, and Wins

An appropriate model for determining the relationship between SLOB and salaries is to ask what determines MLB salaries. While we surmise the answer to include years in MLB, age of player, position of player, and SLOB, we use simple models of SLOB and salary metrics. We suspect that first-year salary will be a function of some collective opinion, which we are not concerned with for this study. It follows that one could compare how an offensive player's salary changes over time relative to SLOB. The scope of this paper is to look at SLOB, salaries, and the World Series.

Salary Research

Heyward and Patrick (2008) found a strong correlation between wins and the revenue and earnings before interest taxes, depreciation, and amortization, EBITDA, of a team. They obtain financial data from Forbes.com. They also find the highest correlation with OBP and wins. Jewell and Molina (2004) use runs and OBP to predict win percentage. They find that the inequity in salaries has a negative impact on winning. Richards and Gruell (1998) study hiring behavior. They use a Probit model to predict the World Series champion. The dependent variables are the mean and variance of the teams' salaries. While the mean salary was statically different from zero, the variance was not.

Salary of a Team

Discussion of salary begs the "chicken and the egg" question, i.e. do the Yankee's have a high payroll because they have good players, or do they have good players because they are willing to pay high salaries. We believe the answers, of course, are yes. We hypothesize that salary, changes in salary, and the percentage change in salary will all have a positive effect on SLOB. Teams pay and make adjustments in pay to obtain players with high SLOB.

We look at the sample of World Series games. We apply Occam's razor to analyze the strength and quality of the most common and straightforward hitting measures. We use the Wilcoxon Signed Rank Sum test to determine the best explanatory power of the models. We look at salary metrics to predict SLOB, and a Logit model to predict World Series champion.

Analysis

Measures

We present the simplest and easiest measure of a pitchers or hitter's marginal productivity. These are Batting Average (BA), Walks plus Hits per Inning Pitched (WHIP), On Base Plus Slugging (OPS), and Slugging times On Base (SLOB).

Methodology

We test the hitting coefficients versus runs as our measure of the marginal product of major league players. We do a multiple regression analysis on the entire sample. There we include indicator variables for wins and the home team, and an index variable for the year. In keeping with Occam's razor, we do a bivariate regression with no constant. The data source is Baseball-Reference.com (2011). We use the most recent data; the World Series Box Scores from 2002 to 2011 (10 series, 53 games, 106 records). Earlier periods do not contain some of the measures, e.g., HPB and SF; these were in neither the boxscore nor recaps, per se.

The estimated multivariate equation is

$$\hat{\boldsymbol{R}} = \boldsymbol{mX} \qquad 8$$

where $\hat{\boldsymbol{R}}$ is a vector of fitted values of Runs-per-game, $\boldsymbol{m}$ is a vector of coefficients, $\mathbf{x}$ is a vector of the constant, a trend variable, indicator variables, WIN and HOME, and one of either BA, WHIP, OPS, or SLOB. The bivariate equation is

$$\hat{R} = mX \qquad 9$$

where $\hat{R}$ is a fitted value of Runs-per-game, m is the linear slope, X is either BA, WHIP, OPS, or SLOB. We look at the variables separately to examine the best fit.

We analyze the economics of buying SLOB. We look at the salary, change in salary from the previous season, and the percentage change in salary from the previous season. The model is

$$\widehat{SLOB} = mS \qquad 10$$

where $\widehat{SLOB}$ is a fitted value of the total SLOB for the team in the World Series, m is the linear slope, and S is either salary, change in salary, or percentage change in salary. We look at these independently to avoid multicollinearity.

We examine predicting World Series Winners. The team's average slob for a series and the percentage change in salaries have the strongest correlations to our World Series wins. Thus,

$$WSC = \frac{1}{1 + e^{-z}} \qquad 11$$

where z is a function of a constant, the percentage change in salaries, and the average SLOB. WSC is an indicator variable that takes on a value of one (winning the World Series) or zero (losing the World Series).

Results

We compute the estimation for the four metrics versus runs in multivariate and bivariate regression models. These look at the productivity. We estimate the resource payment models via linear regression. These look at the payment of that productivity. Then, we predict the World Series winner using a Logit model. This looks at productivity and payments.

Productivity: Multivariate Models

The baseball coefficients were statistically different from zero at the 1% level in all the models. The constant coefficients except for Model 4 (SLOB) were statistically different from zero at the 1% level. The intercept for Model 4 was significant at the 5% level. The intercept is the effect on runs of losing on the road.

Table 1: Multivariate Regression Results

	Model 1		*Model 2*	
Variable	***Coefficient***	***T-stat***	***Coefficient***	***T-stat***
Constant	-5.038	-5.809*	-2.821	-4.204*
Win	0.802	1.776**	0.526	1.189
Home	0.100	0.251	-0.457	-1.184
BA	34.400	11.095*		
WHIP			5.103	11.897*
Year	0.120	1.837**	0.050	0.811
Adj. R^2	0.636		0.664	
F	46.860*		52.799*	
Panel B				
	Model 3		*Model 4*	
Variable	***Coefficient***	***T-stat***	***Coefficient***	***T-stat***
Constant	-5.019	-6.739*	-0.995	-2.097**
Win	0.487	1.181	0.425	1.136
Home	-0.187	-0.519	-0.170	-0.518
OPS	12.601	13.196*		
SLOB			37.187	15.140*
Year	0.051	0.883	0.056	1.063
Adj. R^2	0.704		0.753	
F	63.281*		81.010*	

* significant at the 1% level, ** significant at the 5% level

For the BA model, the Win indicator and trend are both significant at the 5% level. Batting averages are better in a win and have improved over time. The coefficients for home field are not statistically different from zero. The bases on ball related coefficients (WHIP, OPS, and SLOB) have negative signs for home field. Home field teams scored less relative to changes other than their batting average.

The SLOB model had the largest adjusted R^2 and F-statistic, 0.753 and 81.01, respectively. The rank order was SLOB, OPS, WHIP, and then BA. All models were statistically significant at the 1% level.

Productivity: Bivariate Models

We estimate the regressions without a constant. 2. See Table 2 in the Appendix. The bivariate slope coefficients were all statistically different from zero at the 1% level. Since the number of games did not indicate a higher or lower coefficient, this suggests independence between series years. The range in R^2s was between 0.775 (BA in 2011) and 0.971 (SLOB in 2002). The average R^2s were 0.833, 0.87, 0.853, and 0.909 for BA, WHIP, OPS, and SLOB, respectively. For seven of the last ten series, the R^2 for SLOB was between 0.871 and 0.919. For six of those series, the R^2 for SLOB was between 0.898 and 0.919. In 2006 (St. Louis versus Detroit), the series had low batting averages (the BA for both teams was 0.2116 versus 0.2432 for all 10 series), three homeruns, four triples, nineteen doubles, and ten walks (base on balls) in five games. Thus, we had OPS with the best fit over WHIP, BA, and SLOB. This could be the case where the additive (OPS) has a higher correlation than the multiplicative (SLOB).

In 2004, there were 20 doubles, 2 triples, 6 homeruns, and 36 walks. WHIP had more explanatory power than SLOB; the WHIP average was 1.426 versus 1.362 for the 10 series. The OPS was less than the average for the 10 series. In 2010 and 2007, the WHIP model had higher R^2s than BA and OPS. In 2010, the average walks per team, per game was 4.79. The hits per team, per game was 9.5 This compares to 8.30 hits per team per game and 3.6 walks per team per game for all 10 years. In the 2007 series, Boston had two games with double-digit scores and Colorado averaged 2.5 runs in its four losses. In both 2007 and 2010, the SLOB model had the highest explanatory power.

Ranking

We hypothesize that the rankings are better for SLOB than those for OPS. We ranked the productivity models by R^2 and performed a Wilcoxon Signed Rank Sum Test. 3. The test statistic is significant at the 1% level.

Payment: Bivariate Results

We observe that payroll affects SLOB. In Table 3, we use salary metrics to predict SLOB. In all models, the coefficients are statistically different from zero. The strongest fit was for Model 5, Salaries, with an F-statistic of 85.373 and an R^2 of 0.818.

For these World Series, we observe an average salary of $92.56 million with a median of $84.87 million. The average change in salaries is $11.08 million with a median change of $13.26 million. The average percentage change was 15.47% with a median of 15.14%. In Model 5, A 10 million-dollar increase in salary changes the World Series SLOB by 70 points (2.63 runs). In model 6, a change in salaries of 10 million dollars from the previous season increased SLOB by over 330 points (12.42 runs). In model 7, a 10% change in salary results in 136.8 more points of SLOB (5.15 runs).

Table 3: Team Payrolls predicting SLOB

	Model 5		*Model 6*		*Model 7*	
Variable	***Coefficient***	***t-stat***	***Coefficient***	***t-stat***	***Coefficient***	***t-stat***
Salaries	0.007	9.240*				
Δ Salaries			0.034	4.738*		
%Δ Salaries					1.368	2.758**
R^2	0.818		0.542		0.286	
F	85.373*		22.445*		7.609*	

* significant at the 1% level, ** significant at the 5% level

Productivity and Payment: World Series Champion

The coefficients are statistically significant at the 5% level. The joint test statistic is statistically significant at the 1% level. The model predicts 15 of 20 results correctly. In conjunction with an average percent change in salaries of only 6.693% for WS Champions and 24.25% for losers, the effect of a percentage change is negative. In other words, the more the percentage adjustment in salaries, the lower the chance of winning. Teams making larger relative changes in player salaries may get to the World Series, but that will not likely result in the championship. This may be a choice of adding one or more players, salary renegotiations, or both. Either way, tweaking the lineup did not pay off most (75%) of the time. A higher average SLOB adds to the likelihood of a championship.

Table 4: Predicting The World Series Winner

	Logit	
Variable	***Coefficient***	***Z***
Intercept	-6.462	-2.357**
%Δ Salaries	-10.810	-1.994**
SLOB	5.804	2.451**
Psuedo R^2	0.389	
χ^2	10.785*	

* significant at the 1% level,** significant at the 5% level,***significant at the 10% level

Conclusions

Our conjecture was that SLOB was a superior metric to measure runs, and runs win games. Further, teams pay for SLOB, and both SLOB and salaries produce champions. With our models and testing, we suggest that SLOB is superior to predict Runs per game. What Pitchers allow and batter-runners earn is SLOB. Runs acquired are a mirror of Runs allowed in a game. The higher is the SLOB, the higher the amount of runs. We prefer the simplicity and transparency of SLOB. Teams pay and adjust their salaries for SLOB. Winning the World Series is the result of SLOB and tuning salary.

The investigators suggest an examination of the historical record by box score for the World Series and Regular seasons. We are addressing further research on the question of ballpark effect. Is there a difference due to the ballpark? Several forecast models are in use by the authors. On a historical note, Babe Ruth, producing the greatest SLOB of his time, was traded to the New York Yankees from the Boston Red Sox. During his time with the Yankees, they won several championships, alas, Boston never did. Lastly, Boston's championship teams in 2004 and 2007 had the second highest salaries next to the Yankees.

Acknowledgement

We thank Robert Bisking and Gabreille Villafuerte for help with the data and our colleagues for helpful comments. We thank Audrey Crowe for her support.

Notes

1. Henry Chadwick is commonly given credit to the creation of BA.
2. We found that six, five, seven, and six out of ten intercepts for BA, WHIP, OPS, and SLOB, respectively were not statistically different from zero at the 5% level. Additionally, with the constant included the SLOB R^2 outranked OPS R^2 (the Wilcoxon signed rank sum Z was statistically significant at the 1% level).
3. The frequency distribution of the ranks does not appear to be normally distributed.
4. Total Player Rating = ABR+FR+BSR is quite complex.

$$\text{ABR} = 0.47\times\text{singles} + 0.78\times\text{doubles} + 1.09\times\text{triples} + 1.4\times\text{homeruns} + .33(\text{bases on balls} + \text{hit by pitch}) - 0.25(\text{at bats} - \text{hits}) - 0.5\cdot\text{Outs on base} \quad (6a)$$

$$\text{FR} = \text{Player Fielding Runs} - \text{League Average Fielding Runs for the position}(\text{team put outs} - \text{team strike outs}) \times \frac{\text{innings played for the player}}{\text{innings played for team}} \quad (6b)$$

$$\text{Player Fielding Runs} = 0.20(\text{put outs} + 2\times\text{assists} - \text{errors} + \text{double plays}) \quad (6c)$$

$$\text{BSR} = 0.3\times\text{steals} - 0.6\times\text{caught stealing}. \quad (6d)$$

References

Baski, M. (Executive Producer), & B. Miller (Director). 2011. *Moneyball* [Motion picture]. United States: Columbia Pictures.

Boswell, Thomas. 1981."Total Average." *Inside Sports*.

Davenport, D.S. 1969. "Collusive Competition in Major League Baseball, Its Theory and Institutional Development." *American Economists 14*: 6-30.

Ignatin, George and Ronald Crowe. 2005."Occam on Deck." Working paper. Presented at the 2005 Southwest Economics Conference.

James, Bill. 1984. *The Bill James Baseball Abstract.* New York: Ballantine Books.

Jewell, Todd R., and David J. Molina. 2004. "Productive Efficiency and Salary Distribution: The Case of US Major League Baseball." *Scottish Journal of Political Economy 51:* 127-142.

Heyward, Peter and Thomas Patrick. 2008. "How Good is a Baseball Owner's Memory? The Importance of Career Statistics vs. Recent Performance in Salary Negotiations." *Proceedings of the Northeast Business & Economics Association:* 5-13.

Lahman, Sean. 2004. Database 5.2. *The Baseball Archive*. Retrieved from Basball1.com.

Lewis, Michael. 2003. *Moneyball: The Art of Winning an Unfair Game*. New York: W.W. Norton.

Richards, Donald G. and Robert C. Guell. 1998. "Baseball Success and the Structure of Salaries." *Applied Economic Letters 5:* 291-297.

Scully, G.W. 1974. "Pay and Performance in Major League Baseball." *The American Economic Review 64:* 915-930.

Summers, Paul and Noel Quinton. 1982. "Pay and Performance in Major League Baseball: The Case of the First Family of Free Agents." *The Journal of Human Resources 17:* 426-436.

Thorn, John (Editor) and Pete Palmer (Editor). 1995. *Total Baseball*. New York: Viking Penguin.

USAToday. 2011. *Baseball Statistics* [Data file]. Retrieved from http://content.usatoday.com/sportsdata/baseball/mlb/statistics.

Woolway, Mark D. 1997. "Using an Empirically Estimated Production Function For Major League Baseball to Examine Worker Disincentives associated with Multi-Year Contracts." *American Economists 41*: 77-83.

Zech, Charles E. 1981. "An Empirical Estimation of a Production Function: The Case of Major League Baseball." *American Economists 25*: 19-23.

Appendix

Table 2: Bivariate Regression Results+

	Dependent Variable					
	BA	***WHIP***	***OPS***	***SLOB***	***N***	***YEAR***
Slope	22.642	3.693	7.615	38.246	14	2011
t-stat	6.686 *	7.012 *	7.29 *	11.677 *		
R^2	0.7747	0.7909	0.8035	0.9130		
Slope	21.725	4.120	7.270	38.654	10	2010
t-stat	6.939 *	7.868 *	6.538 *	10.025 *		
R^2	0.8425	0.8731	0.8261	0.9178		
Slope	21.321	3.864	6.775	35.587	12	2009
t-stat	8.603 *	9.861 *	9.54 *	10.927 *		
R^2	0.8706	0.8984	0.8922	0.9156		
Slope	17.025	2.980	5.730	29.153	10	2008
t-stat	6.762 *	7.7 *	8.054 *	10.125 *		
R^2	0.8356	0.8682	0.8782	0.9193		
Slope	20.878	3.567	7.336	37.135	8	2007
t-stat	6.23 *	7.718 *	5.897 *	15.187 *		
R^2	0.8472	0.8948	0.8324	0.9705		
Slope	15.914	2.806	5.518	29.138	10	2006
t-stat	7.212 *	7.796 *	8.04 *	6.306 *		
R^2	0.8525	0.8710	0.8778	0.8154		
Slope	17.746 *	3.398 *	5.967 *	30.264 *	8	2005
t-stat	5.883	6.579	6.329	6.883		
R^2	0.8318	0.8608	0.8512	0.8713		
Slope	21.726	3.490	7.048	34.171	8	2004
t-stat	6.165 *	8.091 *	5.795 *	7.957 *		
R^2	0.8445	0.9034	0.8275	0.9004		
Slope	13.248	2.674	5.094	28.159	12	2003
t-stat	6.287 *	7.701 *	8.141 *	9.834 *		
R^2	0.7823	0.8436	0.8577	0.8979		
Slope	22.095	3.986	7.955	36.047	14	2002
t-stat	8.569 *	10.306 *	9.873 *	20.96 *		
R^2	0.8496	0.8910	0.8823	0.9713		

+ No constant by year, * significant at the 1% level, ** significant at the 5% level

Natural Resource Revenues And Increasing External Debt: Are These Enemies Of Existing And Potential Manufacturing? A Case Study Of Kazakhstan

Sarvar Gurbanov, Qafqaz University, Baku – Azerbaijan
Edward T. Merkel, Troy University, Alabama – USA

Abstract

Huge amounts of natural resource revenues have ended up with an excessive volume of external debt. External debt is not harmful and destructive on its own. But adding excessive external debt to the natural resource revenues accelerates real exchange rate appreciation which makes the manufacturing sector more vulnerable. Natural resources with an excessive external debt are not only reason behind a shrinking manufacturing sector, but are a major obstacle to the potential development of the sector which is the primary object of national industrialization policy. This paper will investigate the causes and consequences of the increasing external debt of Kazakhstan.

Introduction

Compared with other former Soviet Union countries Kazakhstan has inherited a better industrial base and also had been able to provide substantial amounts of grain production (Luong, 2000). At the same time the Kazakhstan economy was highly integrated with the Soviet Union. Because of this basic fact, Kazakhstan's industry was hit by the negative outcomes of the collapse of the U.S.S.R. (Pomfret, 2006). As an important windfall gainoil revenues can play a crucial role. But for the Less Developed Countries (LDCs), oil revenues generally are not managed wisely. Instead of being a source of stable economic growth, oil revenues remained as huge amount of funds which are generally mismanaged. Substantial amounts of fund inflow in terms of dollars to the national economy generally exceed the absorption capacity of national economy. Also, oil revenues cause the real exchange rateto appreciate. Exchange rate apprecation brings up less competetive and less affordable exports. By losing share in international markets national industrial production starts to shrink down. In the long run, this process causes de-industrialisation. De-industrialisation means negative growth rates in non-oil industries such as manufacturing. In Nigeria's case, agriculture was adversely impacted as well. Nigeria was an agricultural products exporter prior to oil windfall gains, but then Nigeria got to the point where it started to be an importer of food (Oyejide, 1986).

Because of severe and dramatic changes due to the transition period, Kazakhstan did not possess a settled industrial base and it is very important to analyse whether oil revenues prevented this process or not. A very basic tool for managing a substantial amount of oil revenues is to establish a national oil fund. Without a national oil fund, oil revenues will be consumed and invested excessively by the public authorities. It resembles a flood without any dam. To curb and mitigate the negative effects of huge inflow of oil revenues, a national oil fund plays a crucial role. Assets of national oil funds are invested abroad or accumulated for the next generation. But as oil fund revenues were mainly used as collateral for external borrowing, real exchange rate appreciation becomes faster and brings more destructive effects. Paying external debt mitigates the negative effects of windfall gains. Vice versa is valid also. As an additional pressure to the real exchange rate, with the oil revenues if there is excessive external borrowing it will make the negative effects severe. The Nigerian, Mexican and Indonesian cases brought a new term to the economics lexicon; boom based borrowing capacity. Indonesia floated 10 billion US dollars of short run loans and Mexican public expenditures grew faster than oil revenues (Usui, 1997). Between 1979 and 1998, Nigeria's public debt stock went from 2 billion US dollars to 30 billion US dollars (Ogunleye, 2008).

As long as these countries do not have a competitive industrial base which exports goods and services to international markets, it means that the same threat prevails for Kazakhstan also. According to one estimate, between 2004 and 2049 a 270 billion US dollars inflow is forecasted to take place in the Kazakhstan economy due to oil revenues. With a 5% discount rate, the net present value of this amount will correspond to the 99 billion USD, or 6600 US dollars per capita (IMF, 2004). Also, Kazakhstan's gross external debt rose 51.3% to $119.2 billion by the end of 2010, compared to $113.2 billion at the end of 2009, as reported by the National Bank.Short-term debt decreased 12.6% to $9 billion, but long-term debt increased 7% to $110.2 billion.The share of the public and publicly guaranteed debt in the total gross external debt volume rose 4.3% to $5.1 billion by the end of 2010, against 3.3% ($3.7 billion) at the beginning of the year. Kazakhstan's gross foreign debt amounted to $124.1 billion by the end of June 2011 (Interfax, 2011). Total external debt in 2008, 2009 and 2010 is 79.8%, 99.7% and 85.2% of the GDP. Public and publicly guaranteed external debt for the same years are 1.2%, 1.5% and 1.2% of GDP. In 2008, 2009 and 2010 oil and gas exports totaled to 43.5, 26.2, and 37 billion USD dollars (IMF, 2011). At first sight it is

very barely seen that the distinctive feature of Kazakhstan's excessive external borrowing is mainly a result of private sector transactions. Also, with the combined effect of external borrowing and oil export revenues, there will be additional pressure on real exchange rate appreciation. In the Kazakhstan example, excessive external borrowing made non-tradable sectors such as construction and real estate expand. Workers in the oil sector pull monthly wages up and cause downward wage rigidity on whole economy, and all these put extra pressure on tradable sectors such as manufacturing and agriculture. In this study, we will analyse the ultimate impacts on the manufacturing sector.

Manufacturing in Kazakhstan

Even though Table 1 provides raw data, it is quite helpful to draw conclusions about Kazakhstan's manufacturing pattern. As the length of data is not proper for adequate time series analysis which would let us run an econometric model, here we find it fruitful to examine the situation with data expressed in current prices. Also, the data given below at current prices does not cloud any policy implications. Comparing 2010 and 2006, the main production increase in industry came from increasing oil production. In Table 1, the mining sector shows this clear trend. In 2009, by analyzing the structure of industrial production shares in gross value added, 58.4% (2005, 53%) is from mining and 35.7% (2005, 40.4%) is from manufacturing. Taking a deeper look at manufacturing, the 35.7% figure is consists mainly of the manufacturing of food products (7%), the manufacturing of petroleum product,the processing of nuclear materials (4.8%) and the manufacturing of basic metals and fabricated metal products (14.5%) (ASRK, 2011). Domestic manufacturing in Kazakhstan is negligible.

Table 1: Manufacturing and mining sectors in Kazakhstan

	Volume of industrial production, mln. tenge at current prices	
	2006	**2010**
Total Industry	6 509 896	12 105 526
Mining	3 761 259	7 419 550
Manufacturing	2 406 501	3 844 659
	Employment, thsnd.	
Total Employment	7403,5	8114, 2
Mining	186,8	193,7
Manufacturing	554,8	565, 6

Source : Agency on Statistics of the Republic of the Kazakhstan [ASRK] (2010). *Kazakhstan in 2010*, pgs. 56, 169.

Employment shares of these two sectors also tell much. Even though the petroleum industry is highly capital intensive (Karl, 1997), between 2006 and 2010, employment in the mining sectorshowed a slight increase, 6.9 thousand. However the manufacturing sector is definitely less capital intensive compared with the oil industry and from 2006 to 2010 employment in this sector only increased by 10.8 thousand. Taking into account of the enclave nature of oil industry, it is clear that manufacturing sector employment expands very slowly. Combining the gross value added contribution of the manufacturing sector with overall industrial production, productivity increases in domestic manufacturing is also unlikely. De-industrialisation means negative growth rates in manufacturing. A slower growth rate in manufacturing is accepted as inevitable in the Russian example (Oomes&Kalcheva, 2007). It means that manufacturing records lower growth rates than other sectors of the economy. Another reason for the expanding mining sector is uranium production. Kazakhstan produces a large share of uranium from mines, 33% of world supply. In 2003, Kazakhstan uranium production was only 3300 tones; in 2010, this figure reached 17803 tones (WNA, 2011).

The current situation of domestic manufacturing stems from the attractiveness of booming oil revenues. So far, Kazakhstan has delayed setting a national industry base as a priority. Kazakhstan desires to double its oil production and become one of the world`s top ten oil exporters. The Kashagan oil field holds a reasonable hope for this purpose. A production sharing agreement was signed in 1997, and Kazakhstan's project budget in 2007 increased from US$57 billion to $US136 billion (Nurkamov, 2010). Still commercial operations are not able to be started in the Kashagan oil field. This oilfield will pump its first oil by the end of 2012 or early 2013 at the earliest(Petroleum Economist, 2011).Given the huge costs of Kashagan, 15 years after the initial agreement it is expected to generate commercial returns. Thus the existing booming oil sector and the negative effects of the transition period leave manufacturing almost vulnerable.

Furthermore, extensive FDI inflow towards the mining sector created upward pressure on wages. Previous studies suggest that wage increases in the domestic labor market of Kazakhstan are relatively higher than other transition economics due to a booming oil sector (Auty 2001). Also, at the same timesmall domestic producers and traders may be suffering from price competitive imports from Russia and China (Global agenda, 2009). In the Kazakhstan economy, major foreign investment in the oil sector helped fuel strong GDP growth between 2000 and 2007, averaging about 10 percent a year. Kazakhstani banks borrowed heavily from abroad to fund a rapid expansion of credit, largely concentrated in construction and real estate. When

the global financial crisis hit and capital stopped flowing into the country, credit growth ground to a halt and property prices slumped (IMF Survey Magazine, 2011).It simply means thatafter the initial period of transition, the manufacturing sector domestically suffered from increasing domestic costs and strong national currency pressure on international markets.

Recent article published in New York Times provides helpful hints about manufacturing sector`s crucial role in economic development: "*Manufacturing's muscle helped make the United States a world power, but its contribution to national income is dwindling. Manufacturing's contribution to gross domestic product — roughly equivalent to national income — has declined to just 11.7 percent last year from as much as 28 percent in the 1950s, according to the Bureau of Economic Analysis. The bureau says that, in 2009, Chinese manufacturers generated $1.7 trillion of "value added," versus America's $1.6 trillion.If you let manufacturing go, over time that will have a negative gravitational pull on innovation. And is the financial sector, which rose as manufacturing declined, an adequate substitute?*"(NYT, 2011). This question brings us to the point, where Kazakhstan financial sector is required to be analysed.

Causes and consequences of rising external debt

As stated above, Kazakhstan banks heavily borrowed from abroad to fuel credit on construction and real estate sectors. When the 2008-9 financial crunch came this process started to be painful. The share of non-performing loans in the portfolios of Kazakhstan banks continued to grow, reflecting the cumulative effects of the recession, real estate price declines and depreciation of the Tenge in February 2009. Overall, the foreign debt position of the Kazakhstan banking sector has been US$ 46 billion in mid-2007 (World Bank, 2010). Unfortunately, recently the main crucial issue discussed widely about Kazakhstan's economy is non-performing loans (NPLs). This is the negative effect of the global the financial crisis on Kazakhstan's economy. When prices in the real estate market plummeted it triggered the volume of non-performing loans to increase. Kazakhstan then experienced a type of dilemma which is mainly seen in western developed countries. To face a developed country crisis with LDC experience leaves the Kazakhstan economy in a vulnerable situation. Facing financial crisis with an unsettled manufacturing sector is definitely worse than having a strong real sector which can make up for the losses in a financial sector in the long run.

A booming construction industry along with booming oil sector reflects the general view of resource rich countries. As huge amount of capital flow into the domestic markets, with the rising oil revenues demand gets much stronger as compared with the pre-boom period. Economic literature makes a distinction between tradable and non-tradable sectors. The prices of tradable goods are determined under the international supply and demand conditions. But in the non-tradablesectors' side, domestic supply and demand conditions function. Strong demand comes with windfall gainswhich makes the demand side put pressure upward on prices of non-tradable goods and servicesin domestic markets. Construction and real estate are two major sectors which can be taken into account as non-tradables. Excessive external borrowing from abroad caused a non-traded sector boom in Kazakhstan. As local banks had relied on external borrowing to finance large increases in construction, real estate and personal lending since late 2007, with the global financial crunch the banks faced an increasingly serious liquidity crisis that in effect ended new lending, deflating the property boom and leaving many projects unfinished. Moreover, many consumers found themselves in indebted. (ADB, 2010). The government launched an anti-crisis plan in the late 2008. A total of $10 billion or 9.5% of GDP, largely from National Fund of the Republic of Kazakhstan, the national oil fund, was pledged. The central bank cut the refinancing rate 9 times in 2009 and lowered reserve requirements. The government also pushed up the social outlays, including a 25% increase in public servants` salaries and pensions (ADB, 2010). Despite some progess in banks` restructuring, non- performing loans still amount to a third of the total and about one-quarter of banks are unprofitable (ADB, 2011).

It seems that a painful process of financial crisis triggered grave action plans for the Kazakhstan economy. For industrial development and industrialization, government plans to implement 162 projects totaling $45 billion in investments during 2010 – 2014 (ADB, 2010).

Policy Implications

So far the manufacturing sector in Kazakhstan is not able to secure an overall priority on economic development. To gain a growing manufacturing sector requires a long period of strategy and work. For example, India singled out electronics, including IT, as a key industry for its development strategy almost 40 years ago. For a long time period, India took steps patiently: the creation of a national network of science and technology institutions; subsidies to private IT companies to help develop and diffuse technologies and support for the building of technical skills in higher education; and minimal to no governmental interference or regulation. On improving complex manufacturing exports, for South Korea it took 40 years to build a national industry. Malaysia Taiwan and China are other success stories and cases of building a national industry (Gallagher & Zarsky, 2007). Thus creating a domestic manufacturing base requires a long time. Kazakhstan should seek the

ways of drawing FDI to non-oil tradable sectors, especially for technology transfers. The best way allocating windfall gains efficiently is to channel them to industries for which a nation enjoys a comparative advantage. Without any exact national strategy, idle resource revenues in Kazakhstan either puts pressure on the exchange rate or fuels a real estate bubble.

For the short run, devaluation can be seen as a possible remedy. Resource rich country examples show that real exchange rate appreciation is somewhat inevitable (Wakemann & Linn et al., 2004). However the continuous inflow of windfall gains will leave devaluation groundless. Again, for the long run, devaluation strategy should be avoided.

Maybe the Kazakhstan economic authorities did not use oil fund assets as collateral for excessive external borrowing, but private sector operations created the similar result. The National Fund of Republic of Kazakhstan (NFRK) is supposed to function as a cushion for the huge magnitude of oil revenues which exceeds absorption capacity in the domestic economy. Oil funds are powerful institutions which can eliminate the budget and current account deficit constraints which let LDC governments feel comfortable during liquidity bottlenecks. As figures provided above show, because of private sector decisions, government is forced to use oil fund resources. Boom based borrowing capacity took place indirectly in Kazakhstan economy. To put it differently, the ultimate outcome of external borrowing of the private sector brought up similar results as if the public authority did it.

Conclusion

One more time it is very fair to observe that these natural resources are not harmful in themselves.The main negative outcome stems from the the mismanagement of windfall gains. The main distinctive feature of the Kazakhstan experience is that booming constrution and real estate sectors are the results of the direct actions of the private sector. Generally, mismanagement of oil revenues in LDCs is related to public expenditure programs. Boom based borrowing capacity functioned because of governments` decisions. But in the Kazakhstan case, excessive external borrowing is not direct result of decisions of the economic authority. What falls on the shoulders of Kazakhstan economic authority on this issue is to reinforce banking sector supervision for preventing worsening situation.

A flawed banking industry brought two negative outcomes: (a) the banking sector as a financial intermediary was not able to channel funds from financial to real sector, and (b) the banking sector mainly fueled the expansion of non-tradable goods and services, that is, construction and real estate. Why was the banking industry that much imbalanced? It is the topic of further research.

To conclude, mismanagement of natural resource revenues and external borrowing can have negative impacts on the existing or potential manufacturing sectors. The Kazakhstan case provides a distinctive sample for a flawed private sector which can be responsible for this kind of result.

References

Agency on Statistics of the Republic of the Kazakhstan [ASRK] (2010). *Kazakhstan in 2010.*

Agency on Statistics of the Republic of the Kazakhstan [ASRK] (2011). *National Accounts of the Republic of Kazakhstan, 2005 – 2009.*

Asian Development Bank [ADB] (2010). *Asian Development Outlook 2010.*

Asian Development Bank [ADB](2011). *Asian Development Outlook 2011 Update, Preparing for Demographic Transition.*

"Devaluation dilemmas; Kazakhstan appears poised to devalue its currency; Kazakhstan may be ready to devalue its currency." *Global Agenda*, 26 Jan. 2009.

International Monetary Fund [IMF]. 2004. *Republic Of Kazakhstan: Selected issues, IMF country report No. 04/362*, Washington D.C: 27.

"Is Manufacturing Falling Off the Radar?", (September 10, 2011), The New York Times. http://www.nytimes.com/2011/09/11/business/is-manufacturing-falling-off-the-us-radar-screen.html?_r=2&pagewanted=2&ref=business (01.11.2012)

"Kashagan sets a date." *Petroleum Economist,* June 2011.

"Kazakhstan on Road to Recovery, But Banking System Still Weak". IMF Survey Magazine: Countries&Regions. (August 17, 2010)http://www.imf.org/external/pubs/ft/survey/so/2010/CAR081710A.htm (01.11.2012).

Karl, Terry L. 1997. *The paradox of plenty: oil booms and petro-states*. London: University of California Press.

Luong, P. J. 2000. "Kazakhstan: The Long Terms Costs Of Short Term Gains" In *Energy and conflict in Central Asia and Caucasus,* edited by Robert Ebel and Rajan Menon. Lanham: Roman & Littlefield Publishers Inc.

Nurkamov, Adil. 2010. "Resource Nationalism In Kazakhstan`S Petroleum Sector: Curse Or Blessing?" In *Caspian energy politics, Azerbaijan, Kazakstan and Turkmenistan",* edited by I. Overland, H. Kjaernet, A. Kendall-Taylor, London and New York: Routledge.

Ogunleye, E.K. 2008. "Natural Resource Abundance In Nigeria: From Dependence To Development". *Resources Policy,*33 (3): 168 – 174.

Oomes, Nienke. and Kalcheva, Katherina. 2007. Diagnosing Dutch Disease: does Russia have the symptoms?.http://www.imf.org/external/pubs/ft/wp/2007/wp07102.pdf *IMF Working Paper,* WP/07/102.(01.10.2011).

Oyejide., T.A. 1986. "The Effects Of Trade And Exchange Rate Policies On Agriculture In Nigeria". *Research Report.* Washington DC: International Food Policy Research Institute.

Pomfret, Richard. 2006. *The Central Asian Economies Cince Independence.* Princeton and Oxford: Princeton University Press.

Usui, Norio. 1997. "Dutch Disease And Policy Adjustments To The Oil Boom: A Comparative Study Of Indonesia And Mexico". *Resources Policy,* 23 (4): 151 – 162.

World Bank [WB], *Country Brief 2010.* http://www.worldbank.org.kz/WBSITE/EXTERNAL/COUNTRIES/ECAEXT/KAZAKHSTANEXTN/0,,contentMDK:20629270~menuPK:361877~pagePK:141137~piPK:141127~theSitePK:361869,00.html(01.11.2012).

World Nuclear Association [WNA] (2011). *World Uranium Mining. http://www.world-nuclear.org/info/default.aspx?id=430&terms=Kazakhstan (01.11.2012)*

An Evaluation Of The Perception Of Study Group Effectiveness The Empirical Study Of Southern Wesleyan University Study Groups

Dr. Miren Ivankovic, Anderson University

Abstract

We have constructed a survey, one for faculty and one for students, and by using Likert Scale, we measured the significance of the answers on several different perceptions involving study groups. Students and faculty do have different perceptions of study groups. Faculty members feel that study groups are more important when it comes to active learning than students do. There are differences between what female students/faculty and what male students/faculty thinks about them. Female survey participants place higher value on the study group activities than do their male counterparts. Study groups do affect students' retention in a positive way. Both groups have a positive perception of the correlation between the study groups and learning. Study groups do help in improving critical thinking skills, do not help with organization and time management skills and they do not improve soft skills.

Introduction

Southern Wesleyan University (SWU) offers non-traditional college degree programs in most major cities in South Carolina. Study groups are an integral part of the university's goal of "student centered" or "active learning" process of education. Students meet with their faculty once a week for four hours and meet as study groups for another four hours independently. An evaluation of the study groups, therefore, is extremely important in order to assess them as a valuable source of the learning process.

The Statement of Purpose

The validity of study groups in academia has been discussed for a number of years now. In many cases employers require exceptional teamwork skills, and study groups provide a great opportunity for the division of labor, partnership, sharing of knowledge and formation of social bonding. They also improve students' communication skills. On the other hand, they could induce a free rider problem, clash of different personalities, decrease in output, lack of individual recognition and more.

The purpose of this study is to determine several issues that relate to study groups: 1) Does the perception of study group effectiveness differ between instructors and students? 2) Does gender determine one's perception of study group effectiveness? 3) Do study groups affect student retention rates? 4) Do study groups affect the learning process in a positive way? 5) Do study groups improve students' critical thinking skills? 6) Do study groups improve organizational and time-management skills? 7) Do study groups improve student's soft skills?
According to Wikipedia, soft skills is a sociological term for a person's "EQ" (Emotional Intelligence Quotient), which refers to the cluster of personality traits, social graces, communication, ability with language, personal habits, friendliness, and optimism that mark each of us in varying degrees. Soft skills complement hard skills (part of a person's IQ), which are the technical requirements of a job. Organizations, particularly those frequently dealing with customers face-to-face, are generally more prosperous if they train their staff to use these skills. Screening or training for personal habits or traits such as dependability and conscientiousness can yield significant return on investment for an organization.

The Hypotheses

Study groups are the central factor of our study. The main purpose of this paper is to empirically test if study groups are effective in delivering educational goals. Our model is trying to predict the relationship between study groups and their influence on retention rates among students, the learning process, critical thinking skills, organizational and time-management skills and students' soft skills. Our model would suggest that there is a positive relationship between study groups and retention rates, learning and critical thinking skills, organizational and time management skills and soft skills. We also predict that the perception toward study groups' effectiveness does not differ between faculty and students, or male and female students.

Research Methodology

A qualitative research methodology will be used in this study through the utilization of a descriptive survey taken by both instructors and students. 570 students and 60 faculty members participated in the survey. The survey was open to randomly selected students that are the members of SWU AGS (adult and graduate studies) undergraduate programs across the State. The entire SWU faculty had the opportunity to participate, and over 60% of them responded to the survey. Due to the nature of collected data (via survey) we converted the answers according to the Likert Scale in order to obtain numerical responses. Most of the answers range from 1 to 4 or 1 to 5. With these types of data and variables, we tested our hypotheses and based our decisions by applying the Chi-square test.

REVIEW OF RELATED LITERATURE

Introduction

Student learning teams are becoming an integral part of college and university instruction. They are one of many activities designed to involve students in their own learning as compared to passively sitting through traditional lectures. The use of one or more of these activities is usually termed "student-centered" or "active" learning. While learning teams are often encountered in traditional programs, they are especially prevalent in evening and weekend programs involving working students. Some uses of the teams are purely voluntary while others are a required part of the academic program. Among these are some which serve as a partial substitute for faculty-led classes. No where are they a more integral part of student learning than among the twenty three partners who subscribe to the services of the Institute for Professional Development (IPD) in Phoenix, Arizona.

The purpose of this research is to analyze the effectiveness of the student teams, both from an instructor and student perspective, on one IPD partner school, Southern Wesleyan University (SWU). SWU offers non-traditional programs in most major cities in South Carolina. The study will provide not only a means of testing the prevalent theories of student-learning but also assessing whether student groups can provide an effective substitute for direct faculty contact. Based on the findings of this research, also included are recommendations to increase student team effectiveness.

Active Learning, the Underpinning for Student Learning Teams

The most important contribution to instruction in the past twenty years has been the accelerated application of active learning. Researchers and teachers have come to understand that a mere lecture is a relatively poor way to support student learning.

Active learning places students at the center of their learning where they are the ones with primary responsibility for their learning. The classroom presentation is only one of many avenues to help students learn factual and conceptual material. Further, through active learning, students have a far greater chance of developing traits important to success in life. These include analytical ability, interpersonal skills, leadership, writing, speaking, and assertiveness.

Examples of active learning are certainly not limited to the topic of this paper. We will be focusing primarily on student learning groups. They include class discussion, the case method, and simulation through activities such as game playing, computer-assisted instruction, laboratory sessions, volunteerism, internships, and independent exploration of anything that a student is studying. The library is an age-old place for active learning.

Dale and Dyland of the University of Wisconsin have an interesting way to demonstrate how active learning, as contrasted to passive learning, increases retention of what students learn. They assert through their research that we tend to remember: 10% of what we read, 20% of what we hear, 30% of what we see, 50% of what we hear and see, 70% of what we say and 90% of what we both say and do (Warren, 1996, pp 39).

Exclusively teaching through the lecture method confines students to learning at a low retention level. Placing students in a more active role moves them to higher levels of retention. As Alexander Astin, one of the most noted scholars in higher education puts it, "The amount of student learning and personal development associated with any educational program is directly proportional to the quality and quantity of student involvement in that program" (Astin, 1993, pp 382).

There are at least two other benefits associated with active learning. First, students usually are happier in universities where active learning is prevalent. They are treated with greater dignity and feel they are in greater control of their life. They may work harder but the feeling of being in charge is exhilarating. Second, students engaged in active learning tend to have greater staying power in college. They are far less likely to become dropouts. They feel as though they belong and do not want to leave the action of learning. Many students feel estranged sitting in lecture settings.

Robert Pace of UCLA is the leading authority on the importance of active learning. He relates students' "quality of effort" to what they learn as a result of their effort. "Quality of Effort" is a measure of student initiative. According to Pace, it is what you don't have to do unless you are taking the initiative. "You don't have to browse in the library, you don't have to make appointments to talk with faculty members, you don't have to make outlines from your class notes, you don't have to go to concerts, you don't have to work on a committee, you don't have to ask someone to read something you have written to see if it is clear, and you don't have to have a serious discussion with students whose personal values are different than yours" (Pace, 1990, pp 115).

Pace has administered many thousands of student questionnaires in universities throughout the country. They are used to correlate student engagement with learning outcomes. Based on his research he arrives at five major conclusions:

Quality of effort impacts student learning gains more than any other learning variable. This means that student initiative is more important than family background, the prestige of the college attended, the percentage of faculty with doctorates and what students have achieved in the past. "Granted the importance of all the elements that influence who goes where to college, once the students get there what counts most is not who they are or where they are but what they do" (Pace, 1990, pp 115).

Grades are important but they are not as accurate a measure of student learning as quality of effort. This is because many if not most tests focus on factual learning. Student initiative helps develop other traits which are not routinely measured on tests.

Learning takes place outside the classroom as well as within. This observation has caused some more progressive universities like George Mason University in Virginia to actually develop co-curricular transcripts where student activities outside of the classroom are recorded along with grades. While George Mason is primarily a university teaching traditional-aged students, such transcripts could easily be developed for non-traditional students.

Faculty and administrators would do well to ensure that their college or university provides many opportunities and ample encouragement for student initiative.

It is important to exert quality of effort in as many diverse activities as possible. Implicit in all of Pace's conclusions is that involvement is truly voluntary. If forced upon students, they will not benefit from it. They will have a feeling that they are depriving themselves of some other activity.

Other research on active learning is also applicable to the rationale for group study as an important aspect of active learning. The work of Ernest Pascarella and Patrick Terenzini come to two major conclusions. First, "time on task" is important. The greater the portion of time in which the student is actually engaged in a learning activity (taking notes, engaging in discussions, answering questions, etc.) the greater the level of content acquisition, Pascarella (98). So active learning is not designed to detract from the learning of facts but, rather, enhance it.

Second, learning material as though you had to teach it to someone else is very effective. Group study affords the opportunity for students to teach one another. According to Pascarella and Terenzini, "The experimental research on peer teaching provides reasonably strong evidence that learning material in order to teach it not only increases student involvement in the process of learning, but also enhances mastery of the material itself, particularly at the conceptual level," (Pascarella, pp 98).

Does active learning have a cost to students? Yes, it is a significant one, and colleges and universities using it cannot assume the students will gravitate to it no matter how many opportunities are offered. Consequently, it is critical that the institution monitor student progress to ensure that students are embracing the idea. The cost is discipline. The discipline involved here is the student taking individual initiative, and it is very difficult for all of us. In his book Awaken the Giant Within, Anthony Robbins speaks of three initiative killers all of us can easily relate to – laziness, fear of change, and force of habit. Force of habit can be especially strong of students coming out a high school where little active learning takes place. These students, who could easily constitute a majority of students entering a non-traditional program, present a special challenge.

An Overview of Student Learning Teams

Student learning teams are part of a broad category of learning opportunities called "collaborative learning." Collaborative learning includes any activity where students work together. Students may have first experienced collaborative learning in a science lab. Perhaps an elementary or high school teacher was forced into it because there were not enough test tubes, laboratory benches, or frogs to be dissected to go around, so students were divided into small groups and asked to collectively take on a task. If it was true collaborative learning, students were probably graded collectively. Later, students encountered it in true science or computer courses where a laboratory experience was an integral part of the course. Of late, "learning communities" have become especially popular on campuses. They focus on helping students see links between courses they are taking. Some may involve a small cohort of students who take several classes characterized by students taking two or more classes which are connected thematically and reinforced by cohort study. Here the faculty do develop and

monitor plans for the cohort study. Finally they may involve student cohorts taking classes which are team taught by the faculty.

In any form, including the SWU student learning teams, collaborative learning can be very important. First, studies have found that most of us reason better in groups. In her book Anatomy of Judgment, Abercrombie showed that a group of medical students collectively diagnosing a patient acquired better medical judgment faster than individual students working alone.

A second reason it is effective is much more obvious. Collaborative learning helps prepare students for the real-world where decisions, indeed work itself, is usually done in groups. That requires learning the set of social skills required to reach effective decisions. David and Roger Johnson summarize the gains of using collaborative learning that causes students to cooperate rather than compete. They contend that more than 323 studies on the subject have been conducted over the past 90 years. Generally, "achievement is higher in cooperative situations than in competitive or individualistic ones and...cooperative efforts result in a more frequent use of higher level reasoning strategies, gains in skills of students working with one another, and gains on class tests given to individual students," (Johnson, pp 91).

Kenneth A Brunfee understands the richness in human relations that can come from collaborative learning. His statement captures its true benefit:

"What we learn in particular is that to work well together, sometimes we may have to give something up. What we learn to give up is always getting to do exactly what we want to do. Or, if we do get to do what we want to do, we may have to give up always doing it in exactly that way we want to do it. In short, people who work well together have learned to share their toys."

"...by learning to share "our toys," I mean learning to share our books, our ideas, our benefits, our way of life, our cities, our country, our world. Most of us spend a lifetime learning to share our toys," (Brunfee, pp 45).

Student learning groups or any form of collaborative learning are not automatically successful. They require work on the part of the students and solid planning and supervision on the part of the college or university. What is required for collaborative learning to be effective? Brunfee helps answer the question. First, each work group needs to have the benefit of some instruction in group dynamics, often called "social psychology." Look for a moment at the questions faced by a new group. How does one find a leader? What are the essential qualities of a leader in this instance? How do you quickly determine the relative strengths of the group members? Which tasks should be assigned as individual tasks? Which tasks should be critiqued? How are conflicts between members of the group to be resolved? What is done about members of the group who do not carry their load?

The last question usually comes up in every group and most guides to collaborative learning suggest that some aspect of the grade has to be based on the contribution of individuals. The problem with that approach is that it requires that the instructor intrude into the operation of the group to determine who is doing what. Perhaps students can grade each other but that intrudes on the mutual dependence of one student on another. Faculty have long wrestled with the problem of evaluating each other for tenure and position and most would admit that doing that conflicts with department harmony. Consequently the evaluations often tend to be overly kind.

The second requirement for successful collaborative learning is that the teacher and then the class members must be willing to grant authority to one another for specific tasks. Easy, you say? In our own work groups some are quite comfortable relinquishing responsibility to others. But still others find it quite unnatural, especially if it means relinquishing power to another.

A third condition is that students must be willing to accept the authority of other students. In other words, there must be effective followers. Students are usually comfortable accepting tasks assigned by an instructor but are usually less comfortable accepting the same task assigned by a fellow student.

The final requirement is that there be friendliness and civility among members. This may come naturally, especially when all is well, but can change quickly when a sense of urgency and pressure are introduced into the environment. Perhaps some of the members come from different cultures and have vastly different perspectives on the assignment. All differences between the individuals have to be accommodated for the good of the project. This means that students have to exercise the skills of negotiation and compromise.

Of course the conditions for successful collaborative learning present not only a challenge but also the basis for student learning. If students can overcome these challenges they are a long way down the path to becoming effective workers in a group and they can carry that skill into their job and even into their home. The importance of that is reinforced by Johnson, Johnson, and Holubes: "Most employers do not expect people to sit in rows and compete with colleagues without interacting with them. The heart of most jobs, especially the higher-paying more interesting jobs, is teamwork, which involves getting others to cooperate, leading others, coping with complex power and influence issues, and helping solve people's problems in working with each other. Teamwork, communication, effective coordination, and divisions of labor characterize most real-life situations," (Johnson, pp91).

Research Design and Methodology

It was determined that the design of this study on the main research problem is adequate to decide if the data used to test the hypotheses is available in a format that grants easy analysis and application. Data collected meet the criteria of establishing a central goal for execution of the research problem. A descriptive survey design has been determined to be the most effective method to collect the data.

Survey

For data collection in this research, a questionnaire survey is the method of selection. Implementing a sequence of questions to a group of instructor and student participants will permit for uncomplicated summarization of responses for doing statistical analysis. For this reason, the chosen method of instrumentation to collect, tabulate and analyze data will be a written survey taken online.

Questionnaire Formulation

The questionnaire design is simple and uses direct language. To reduce bias, wording of questions are clear, brief and concise dismissing the use of infrequent words or long sentences, which eliminates misunderstandings, appearing easier to complete. The instructor survey consists of 19 statements, all presented in a positive form. Two additional questions soliciting free-form comments are added to the end.
The student survey consists of 35 statements, all presented in a positive form. Four additional questions soliciting free-form comments are added to the end.

General Order and Format

Identifying the demographics of each participant is established in the beginning of the survey. Planning, formulating and reviewing the data based upon the three hypotheses in this research, designed the logical structure and flow of remaining statements and assisted with judgment of question order and format. An introduction to the survey briefly states the purpose of the research and ensuring participant responses are confidential and not personalized.

Order of Question Categories

The first six questions on the instructor's survey refer to demographic data. The first fourteen questions on the student's survey are of a demographic nature. Additional survey questions flow in a logically coherent form.

Response Modes and Scales of Measurement

The participants will be asked to classify whether an agreement or disagreement to each statement exists and to what strength. The rating scale item used within the survey to measure this data is the Likert Scale. This is the most common scale given in an ordered or logical relationship of close-ended statements within the survey, establishing reliability. Clarification of responses for each statement is given in numerical value form from one to five.
A sample question is: Overall, study groups are important to my learning.

1. Strongly Disagree 2. Disagree 3. Neither Agree nor Disagree 4. Agree 5. Strongly Agree

Participants

The survey will be made available online to all instructors and students attending class during a specified week in November 2005. Taking this survey will be a mandatory assignment for the students and required from each instructor. The survey will be made available online for both students and instructors. The participants will be able to access the survey from any computer through the internet connecting to the Southern Wesleyan site. After completion of the survey, the instructor or student will be issued a code attesting that they took the survey. All responses will be completely anonymous.

Methodology

A qualitative research methodology will be used in this study through the utilization of a descriptive survey taken by both instructors and students. 570 students and 60 faculty members participated in the survey. Survey was open to randomly selected students that are the members of SWU AGS (adult and graduate studies) undergraduate programs across the State. The entire SWU faculty had the opportunity to participate, and over 60% of faculty responded to the survey.

Due to the nature of collected data (via survey) we converted the answers according to the Likert Scale in order to obtain numerical responses. Most of the answers thus range from 1 to 4 or 1 to 5. With these types of data and variables, we tested our hypotheses and based our decisions by applying the Chi-square test.

Testing the Hypotheses

1) Does a difference exist between instructors and students at Southern Wesleyan University as to the perception of study group effectiveness:
H0: Students and faculty have the same perception of study groups.
H1: Students and faculty have different perceptions of study groups.

Table 1: Chi-square Contingency Table Test for Independence:

	B	G	Total
STUDENTS	279	859	1138
FACULTY	112	65	177
Total	391	924	1315

The values of the Chi-square is 110.14 > CV, df = 1
p-value 9.11E-26

Based on the result, we reject the Ho and conclude that students and faculty do have different perceptions of study groups.

2) Does gender matter in the perceptions of study group effectiveness; the hypothesis is:
H0: Gender does not affect perception of the group.
H1: Gender does affect perception of the group.

Table 2: Chi-square Contingency Table Test for Independence:

	Col 1	Col 2	Total
Row 1	202	78	280
Row 2	657	201	858
Total	859	279	1138

The values of the Chi-square is 2.24 < CV = 3.841, df = 1
p-value = .1345

Critical value (3.841) > 2.24 (chi-sq. score), thus we reject the Null hypothesis and conclude that gender does affect the perception of the study groups. Basically, there are differences between what female students/faculty think of study groups and what male students/faculty thinks.

3) Do study group affect student retention rates; our hypothesis is:
H0: Study groups do not affect students' retention.
H1: Study groups do affect students' retention by reducing the number of students' quitting the program.

Based on the Chi-square, 387.37, (very high), we reject Ho and can conclude that study groups do affect students' retention in a positive way.

4) Do study groups affect learning process in a positive way; our hypothesis here is the following:
H0: Study groups are important to learning.
H1: Study groups are not important to learning.

Based on the Chi–square = 5.87 < CV = 76.78, (df = 58, p-value = 1.000) thus we Fail to reject Ho and conclude that study groups are important for learning activities.

5) Do study groups improve students' critical thinking skills? Hypothesis is:
H0: Study groups help improve critical thinking skills.
H1: Study groups do not help improve critical thinking skills.

Chi-square = 157.21 < CV, (df = 568, p-value = 1.000) thus we fail to reject Ho and conclude that study groups do help in improving critical thinking skills.

6) Do study groups improve organizational and time management skills; the hypothesis is:
H0: Study groups help improve organization and time management skills.
H1: Study groups do not help improve organization and time management skills.

Chi-square = 214.13 > CV, (df = 568, p-value = 1.000) thus we reject the Ho and conclude that study groups do not help with organization and time management skills.

7) Do study groups improve student's soft skills? The hypothesis is that:
H0: Study groups do improve soft skills.
H1: Study groups did not improve soft skills.

Based on our Chi-square of 357.73 and its critical value, (df = 1137, p-value = 1.000) we have to reject the null hypothesis and conclude that study groups do not improve the soft skills.

From the above results, there is a different perception about the purpose of study groups between students and faculty, but variable on gender shows the same perceptions about study groups. Study groups positively affect retention, learning and critical thinking. The perception is that study groups do not improve organization, time management skills or soft skills.

We have also tried to test the notion that study groups improve learning and knowledge. We selected 30 undergraduate business major students from SWU, which were given a Major Field Administration Test (MFAT) in Business before their graduation. We also randomly selected another 30 undergraduate business major students who took the same test, but this group was selected from SWU non-traditional (night) programs where formation of study groups is required throughout the whole program. The test was conducted in the spring of 2005. The first group (traditional students) scored a mean score of 147 with a standard deviation of 5.4, while the second group scored 154 with a standard deviation of 4.8. The scores were statistically different using a t-test for 2 independent samples. National scores that semester were about 153 points. The group that was required to form and maintain study groups outperformed the other group and also performed above national averages. This is yet another indicator that study groups do enhance learning and knowledge.

Conclusion

A large amount of literature and research has been devoted to the concepts of learning, best teaching strategies, classroom activities and other similar topics related to education. Strategies that involve active learning show very promising results when it comes to the absorption of educational materials. In this study, we went a step further and used the study groups as the independent variable and tried to measure students and faculty perceptions about them as a viable source of active learning.

We have constructed a survey, one for faculty and one for students, and by using Likert Scale, we measured the significance of the answers on several different perceptions involving study groups. We have found that students and faculty do have different perceptions of study groups. Going through the responses, faculty members feel that study groups are more important when it comes to active learning than students do. Second, we found that there are differences between what female students/faculty think of study groups and what male students/faculty thinks about them. Female survey participants place higher value on the study group activities than do their male counterparts. Third, we found that study groups do affect students' retention in a positive way. Study groups work like a social cell, and there is a peer factor or a team factor that bonds students. This makes the cost of dropping out of school somewhat higher. Fourth, we found that both groups have a positive perception of the correlation between the study groups and learning. Fifth, study groups do help in improving critical thinking skills. Sixth, study groups do not help with organization and time management skills. And seventh, they do not improve soft skills. The last two findings are a bit surprising, especially the sixth. One would think that the perception of study groups is that they make the participants more organized and punctual since they have to bring to the group their share of the work, which does require both organizational and time management skills.

Overall, we find that students and faculty, in the "eyes" of our sample, value study groups and what study groups are supposed to achieve, as well as provide us with some new research ideas that can add more value to this paper. Active

learning is beneficial, especially in classes where students are not motivated to listen for long periods of time, and new strategies and methods should be applied in order to correct this problem.

Acknowledgements

Faculty would like to express special thank you to the team of research assistance from Anderson University, Heather Martin, Rob Wallace and Jan Schieber.

References

Alexander Astin, 1993, *What Matters in College: Four Critical Years Revisited* (Jossey-Bass), 382.
Beth-Mackowiak Ayotte, Stacey, 2010, "Is There a Place for Games in the College Classroom?" *Tips for Encouraging Student Participation in Classroom Discussions.*
Brunfee, Kenneth, "The Art of Collaborative Learning," *Change Magazine*, April, 4 1987, 45.
Chapnick, Adam, 2010, "Creating a Class Participation Rubric," *Tips for Encouraging Student Participation in Classroom Discussions.*
Cieniewicz, Jon, 2010, "Participation Blues from the Student Perspective", *Tips for Encouraging Student Participation in Classroom Discussions.*
Johnson, David, Johnson, Roger and Johnson, Edythe, 1990, Circles of Learning: *Cooperation in the Classroom*, (Edina, Minnesota: Interaction Book Co., 91.
Knight, Denise, P, 2010, "Assessing Class Participation: One Useful Strategy", *Tips for Encouraging Student Participation in Classroom Discussions.*
Lathrop, Anna, H, 2010, "Teaching How to Question: Participation Rubrics", *Tips for Encouraging Student Participation in Classroom Discussions.*
Pace, Robert, 1990, *The Undergraduates: A Report of their Activities and Progress in College in the 1980s*, UCLA Teaching Center, Los Angeles, 115.
Pascarella, Ernest and Terenzini, Patrick, 1991, *How College Affects Students*, San Francisco: Jossey-Bass Publications, 98.
Skinner, Nicholas, F, 2010, "Discouraging Over Participators", *Tips for Encouraging Student Participation in Classroom Discussions.*
Swope, Kurtis, J, 2010, "Roll the Dice and Students Participate", *Tips for Encouraging Student Participation in Classroom Discussions.*
Taylor, Ann, 2011, "Top 10 reasons students dislike working in groups and why I do it anyway", *Biochemistry and Molecular Biology Education,* Biology Education, 39 (2), 219 – 220.
Warren, Russel G, 1996, *Carpe Diem*, University Press of America.
Weimer, Maryellen, 2010, "Putting the Participation Puzzle Together", *Tips for Encouraging Student Participation in Classroom Discussions.*
Weimer, Maryellen, 2010, "Student Recommendations for Encouraging Participation", *Tips for Encouraging Student Participation in Classroom Discussions.*
Weimer, Maryellen, "Those Students Who Participate Too Much", *Tips for Encouraging Student Participation in Classroom Discussions.*
Weimer, Maryellen, 2010, "To Call On or Not to Call On: That Continues to Be the Question", *Tips for Encouraging Student Participation in Classroom Discussions.*
Weimer, Maryellen, 2011, *Practical Ideas for Improving Student Participation.*
Weimer, Maryellen, 2012a, *My Students Don't Like Group Work.*
Weimer, Maryellen, 2012b, *Questions About Active Learning*, Faculty Focus.

Re-employment Loan Approach to Economic Recovery

Richard Lewin, Marc Sardy, Rollins College

Abstract

We examine the impact of accounting and tax treatments afforded to labor and capital in a corporate context. Labor costs appear disadvantaged with respect to capital investments, due to cash flow benefits from depreciation allowances and interest expense deduction. In terms of international competitiveness, re-employment favoring policies might be justified to place labor inputs on a more equal footing, thus enhancing domestic job creation. We frame our investigation in terms of a re-employment cycle loan for human capital - a portable government-sponsored loan to corporate employers tied to new hires - based on the NPV of future welfare payments.

Introduction

In emerging from the Great Recession, whilst some of the previous media jargon of 'stagnation' has been replaced with more contemporary counterparts of 'jobless recovery' and 'downsizing', the fundamental issue that faces developed nations' government remains the same: how to get people back to work. Policy makers are being pressured into taking another active role in improving employment opportunities and prospects, once again making tax incentives designed to stimulate jobs a real option that is open for discussion.

Many economists believe that capital accumulation, technical progress and labor force expansion have no lasting effect on unemployment. This view rests on the empirically dubious assumption that the elasticity of substitution between labor and capital is equal to unity (i.e. production is a Cobb-Douglas function). Capital accumulation has supply-side effects on the labor market however if the elasticity of substitution between capital and labor is not equal to unity (Rowthorn, 1999a, Arestis and Sawyer, 2005). With a lower elasticity of substitution, the equilibrium unemployment rate is affected by all of the above factors. In agreement with the empirical studies, the elasticity of capital-labor substitution has been shown to be below unity in the literature; for a summary see Chirinko (2008). This we would argue is the pervading issue in most western democracies and is due to a relentless substitution of capital for labor due to the inherent advantages that accrue to capital versus labor on the financial credentials of corporations. The substitution of capital for labor and vice versa, is nothing new; it is at least a 200-year old idea, Hayek has written several important papers on the topic which can be found in a collected volume of his work (Hayek, 1989).

Consideration of the role of improvements in the quality of the labor force (due primarily to education and on-the-job training) helped 'explain' observed economic growth, and stimulated development of the concept and theory of human capital (in this context, human capital may be defined as skills and knowledge that enhance a worker's productivity in the labor market). Note, however, that the idea that human beings may be regarded as analogous to physical capital (where capital may be defined as a produced means of production) can be traced back to Adam Smith (1776).

Becker (1965) may serve as the basic conceptual framework. In general, then, it should be clear that as wages rise, time-intensive commodities become increasingly expensive; and commodities will increasingly be produced via goods-intensive technologies of production in order to economize on the use of time. Capital accumulation may be endogenous and if so what will the long-run implications of this have for unemployment?

In the economics literature, the NAIRU model has become the dominant framework for the macroeconomic analysis of unemployment, as proffered by textbooks such as Blanchard (2006) or Carlin and Soskice (2005). These detail how equipment and software prices are getting rapidly cheaper while the cost of labor has been getting more expensive, making capital a far more attractive investment to companies than people. Tax incentives that encourage earlier capital investment as introduced in recent legislation may ironically be helping to tip the balance too against increased labor demands.

Typically direct labor-market fiscal interventions have been in the form of either marginal or targeted-recruitment tax credits, rebates or subsidies for employers. The former is a credit against tax for any firm that adds incremental employees to its labor force, while the latter is available to firms that hire specific types of workers – usually the 'unskilled' or 'disadvantaged.' Marginal employment tax credits, rebates or subsidies in effect, subsidize firms' wage costs (directly to the firm in the form of credits or rebates, and indirectly in the form of tax subsidies) or indirectly as an incentive for employers to increase production and labor demand, and/or substitute new workers for capital. Importantly this only makes sense if capital and labor are substitutes in production. Naturally, we might think of them as complements. Indeed until capital can do everything labor can do – that is until the singularity condition is reached - assuming a business cannot entirely replace the worker with a CNC welding robot – some types of jobs must necessarily be compliments to capital. Those jobs will always

be in more demand as capital gets cheaper. The question is how much skill you need to do those jobs. This is the whole issue of skill-based technological change, which seems to be polarizing the educational attainments in most developed economies.

Even though we see movement of labor-intensive production from high labor cost countries to low labor cost countries, the subsidies to capital may be accelerating the decline of labor even in those low-cost labor markets. China Business News and the Financial Times newspaper (Hille, 2010) recently reported that Taiwanese electronics manufacturer Foxconn, which assembles products for firms including Apple, is sharply increasing its use of automated equipment in its factories. In 2011 a BBC news service article cited IHS Global Insight analysts in Beijing, adding that Apple was 'leading the way in the next phase of automation for low-cost production'. The move towards automation was aimed at shifting 'workers from more routine tasks to more value-added positions in manufacturing such as research and development, innovation and other areas that are equally important to the success of our operations'.

Fiscal policy may promote higher rates of investment thus raising the level of capacity in the economy, and therefore employment, in the longer run. A country may affect the level of employment either by altering relative output prices to induce consumers to switch to more labor-intensive products, or by altering relative factor prices to induce producers to shift to more labor-using techniques. Relative factor prices in favor of labor within the organized sector can be done as equivalently by making capital more expensive, as making labor cheaper capital - biased tax incentives really can have the effect of discriminating against the employment of labor. Reviewing the related economic literature from the OECD (1995 & 1996), support was found for the following four general observations about tax incentives and job creation.

Effectiveness in increasing the employment of workers who would not have been recruited in the absence of a tax credit is mediocre at best. As a result, these programs generally incur high costs per net new job according to the cases studied in these surveys. Tax incentives are preferred to direct government job creation programs. As Robert Solow (1980) concludes:

'[because] profit incentives operate more or less as they are supposed to ... wage subsidies have some advantage over direct job creation according to the efficiency criterion. They are probably also to be preferred according to the equity criterion: they offer at least the possibility of a start in the mainstream labor market, whereas direct job creation at least runs the risk of creating a sort of caste [system for labor force pariahs]'.

The literature supports the notion that targeted employment tax credits are preferred to marginal employment credits, even though there is no proof that they are any more cost-effective. Nor is it clear that these credits lead to sustainable net job growth. The rationale is one based on economic, political and bureaucratic factors. Firstly, it is easier to measure effects and control abuse for programs that single out particular types of individuals, than for those programs that broadly attempt to assist anyone who is unemployed. Secondly, according to Gera (1988) directly assisting those workers who possess hardly any bargaining power minimizes any inflationary impact on wages. Thirdly, if it comes in the process of helping those who are least able to help themselves, any consequential displacement of non-subsidized workers may be socially acceptable.

Tax incentives for job creation certainly prove to be expensive, although they are not completely ineffective since firms do seem to respond to varying degrees and some jobs will ultimately be created. The dilemma facing government, and by extension the electorate, is in being to determine at what price the cost per job created is actually too high.

The best evidence for the potential of human capital contracts is a government backed system in Australia, where students repay the costs of attending college through forfeiting a percentage of their after-college income. Instead of a fixed time-period, students continue paying until their entire debt with reasonable interest is repaid.

It is proposed that equivalent re-employment loans could be portable and written off against employer/employee taxation over the residual working life, but would be repayable in full should workers be sacked or poached by competing firms. We envision a partnership between both public and private sector; both of whom could benefit from establishing a re-employment bond market, supported by private investors albeit guaranteed by government - in similar fashion to the market for student loans. An incentivized investment in human capital by employee / employer through training and further education could moreover supplement the scale of such loans against enhanced future remuneration, in the same way as student loans are extended at the graduate level. A long-term employment loan, sensitive to life cycle income, education and skills, is thus proposed. We provide an algorithm for deriving such an employment loan offered over the residual employment period with a fluctuating interest rate tied to unemployment levels. Thus, government-sponsored loan facilities could potentially fill the unemployment vacuum left in the wake of the credit crisis and Great Recession through a meaningful expansion of re-employment and training opportunities.

We look at a guaranteed public sector impact for value added tax revenues, where actual money spent puts the onus back on private industry. Currently there is no incentive for employment neither for business nor as employee specific. Employees could be offered a portable life-cycle loan, augmented by length of employment, perhaps for up to 50 years and guaranteed by government to allow companies access to this capital stock. Thereby acting as an incentive to retain good employees and grow their human capital by further investment, which may augment the size of loan available to the business per worker. The incentive to stick around and not incentive job hopping may change the nature of employment and could provide higher social utility. Capital benefits from additional deductions associated with its expenses and with maintaining on-going expenses trade-off to labor and capital cost. As we look at the interest and depreciation tax shields, it should be noted that

there is a marginal benefit to capital. Our aim is to equal the NPV of costs of labor versus investment in capital goods. Proof of economy wide benefit, based on choices between labor and capital are needed. More generally, this might provide a level playing field for investments in human capital, thereby stemming the tide of relentless substitution of capital for labor as witnessed throughout much of the developed world.

Current corporate tax treatment, in almost all western democracies, allow for firms to depreciate the value of their capital equipment over time, ostensibly to ensure a pool of cash is available for future replacement of such equipment. However, effectively it is often available to be used for any purpose that the company could deem necessary. In addition, firms can also enjoy the advantage of an interest tax shield associated with any debt accumulated by purchasing such capital equipment. This combination of benefits, created by favorable tax treatments, generates an effective marginal benefit to capital equipment versus labor, further causing the Cobb-Douglas production function to vary from unity.

Conversely, labor as an input is purely used as a source of expense on corporate income statements and yields neither depreciation benefits, nor any associated interest tax shields. When compared to capital equipment where the only expenses are replacement parts and service, labor has a series of additional implicit incremental costs, timing of which is borne directly by the firm and affects its cash flow; such as the timing of any national social security contributions and the sourced income tax arrangements for employees alongside any pension contributions. All of these costs are often remitted to government on the same timing basis as they are remitted to the employee, rather than in arrears as the remittance of annual corporation tax payable on profits. The nature of this retrospective payment after the year-end ensures that working capital resources stay in the firm longer, enhancing liquidity and allowing interest payments on depreciable capital to be met from this cash source in the interim.

$$\sum_{t=1}^{n} \frac{Annual\ Depreciation * Corporate\ Tax\ Rate}{(1+r)^t}$$

Where: t=period; n=number of time periods; r=interest rate

Annual depreciation allowance allows for the recovery of capital based on the economic life of an asset multiplied by the prevailing corporate tax rate. The PV of the subsequent recovery will be less than dollar based on the useful life of the asset and the cost of capital. For example, under US Modified Accelerated Cost Recovery System (MACRS) given a 8.29% cost of capital assumption, the NPV of a 7 MACRS depreciation schedule amounts to a $0.701 recovery per dollar of initial investment. Furthermore, technology and capital equipment often lead to lower real cost replacement price for more efficient equipment at the time of full depreciation or replacement.

To the extent that the investment is debt-funded the interest tax shield further reduces the effective cost of the investment by lowering operating profits. Debt also extends the leverage of the firm at a multiple of the after-tax interest rate. So if the cost of labor < depreciation expense - depreciation tax shield + interest expense - interest tax shield plus cash flow timing benefits

$$\lambda = \frac{K}{n} - \sum_{t=1}^{n} \frac{\frac{K}{n} * \alpha}{(1+r)^t} + Kd - \sum_{t=1}^{n} \frac{Kd * \alpha}{(1+r)^t} + \varphi$$

Where:

K = Capital equipment
n = Number of depreciation periods
Kd = Cost of Debt
r = Levered Cost of Capital
α = Corporate tax rate
□ = Cash flow timing benefits
λ = Marginal Benefit of Capital

If:

$$\left.\begin{array}{c} Cost\ of\ Labor > \lambda\ replace\ labor\ with\ Capital \\ Cost\ of\ Labor < \lambda\ replace\ Capital\ with\ Labor \\ Labor = \lambda\ parity\ condition \end{array}\right\}$$

The firm should choose to use labor (all things being equal) in preference to capital. Conversely, if the cost of labor is significantly higher as we contend in our paper, then the firm would most often choose to replace labor with capital equipment wherever possible. This is not inconsequential at the aggregate level, when firms can add to the investment in capital equipment in manufacturing for example, which is less expensive versus labor, depending on the market. This can be

much more expensive in the developed world than in terms of expenses in the developing world. Often in the developed world, many firms will look to replace labor with capital or outsource operations internationally where gross labor rates are lower.

Over the years, there have been significant advances in capital equipment and efficiency. Demands for high precision and continuous quality improvement have also been accelerated, replacing labor with capital equipment. Yet, this stands in stark contrast with what is required for an economy to be successful. While productivity enables an economy to improve output, lower cost and increase the marginal benefit to all, capital equipment cannot spend its wealth in the economy to drive the underlying multiplier effect and thus create demand and jobs. It can only create low-cost goods that perhaps extend the apparent wealth of consumers in an economy where disposable income is prevalent. However, if in the economy the labor capital trade-off yields an increasing substitution of capital for labor, unemployment rates will inexorably climb, disposable income increase will decline and ultimately an economy may spiral inward, under the social support network required to sustain the labor force, which may be highly skilled yet largely unemployed.

So how does a country fix the apparent inequity between labor and capital? Invariably the most important factor is to level the playing field for decisions in respect of investments in labor and capital. If we consider labor the equivalent of human capital then there is a substantial literature about the depreciation of human capital (Becker, 1965, Arrazola, et al, 2005). Yet there is no associated depreciation tax expense with this form of capital investment at the individual or corporate level, and thus no depreciation tax shield available either. Furthermore, there is no interest expense associated with this form of capital even though significant sums are spent by firms worldwide in maintaining the currency of employees' knowledge and thus no interest tax shields are available from monies spent in this category of investment aside from appearing as expenses.

It is our premise that in order to reverse the inward spiral of economies which leads to higher unemployment as a result of capital equipment replacing human capital, it is necessary to level the playing field. How might this conceivably work? Consider first the depreciation of capital equipment; it works as a set-aside to replace the capital equipment when it ultimately has outlived its usefulness. The expertise associated with labor also depreciates over time. In order to sustain or improve human capital it is necessary to invest throughout the productive lives of workers. Perhaps it might be possible to develop a set-aside within a firm for investments in human capital. This would have the effect of increasing the knowledge base and employability of members of the workforce. It would ultimately lead to an improvement in human capital and a more skilled and better-trained workforce. An improvement in human capital leads towards a virtuous circle of innovation and creativity, which ultimately can add significant value in terms of comparative advantage to any economy that embraces it. Ultimately, this form of depreciation could act as a set-aside for an educational improvement of the firm's human capital. The net effect of this may ultimately reduce tax revenues for the respective government in the short term although higher employment levels commensurate with higher wages in the longer term may raise the net tax take per unit of investment in labor production.

Yet, consider that in the current global tax environment, competition for businesses with high labor demands has led to spiraling corporate tax rates and more tax harmonization as countries adopt 'beggar-thy-neighbor' policies. It is conceivable that a firm operating within the boundaries of a country that has this form of appreciation tax credit may see that tax benefit as an overall reduction in the corporate tax rate. The practical benefit to the firm is a reduction in the tax rate commensurate with the depreciation tax shield of human capital.

Consider that reemployment bonds would generate financial capital that could be used in times of tight credit. This capital would come at the expense of the investment community with guarantees from the government. This would directly connect investment with reemployment, which would yield more taxable income and higher aggregate consumption. This should lead to a higher aggregate tax for the country from the re-employed population.

Turning our attention to interest tax shields there needs to be a comparable shield for labor as well. Consider that interest tax shields come from extremely large investments in capital equipment that result from debt necessary to finance the purchase of the equipment. The interest tax shield allows firms to leverage more effectively their borrowing and investment practices and ultimately maximize the returns to shareholders. Yet it could be conceivable that a human capital tax shield as a direct result of some form of debt incurred in training or maintaining the existing currency of employees might level the playing field for firms more interested in obtaining overall savings for the firm.

We propose a loan, which by its nature would be directly attached to human capital re-employment. Consider a loan which might be established in the private sector and guaranteed by the public sector. Models already exist for human capital loans in the form of student loans. Yet the burden of repayment of student loans are borne solely by the individual, who has directly benefited from the education in terms of potentially higher wages, but would have no means of reclaiming this investment by way of associated depreciation. In economies where the debt burden associated with education is high, salaries of newly educated human capital are considerably higher than their counterparts in countries where education costs to the individual are lower, or the education level is lower. In other words, the additional costs of the loan are passed on to the firms that hire. This is yet another reason why many firms have moved their labor-intensive industries away from countries where labor costs are even marginally greater still. If these reemployment loans were also given the same tax benefits that debt

associated with capital equipment receives it might reduce the incentive to switch from labor to capital especially in the most developed economies with the highest education levels. Furthermore, if one considers that the loan might be present throughout the lifetime employment of human capital with the firm, where repayment in full was due upon the dismissal or departure of the employee, there would now be a switching cost that might give the firm an even larger disincentive to reduce its human capital and instead retain and enhance its existing workforce.

To recap, a new bond perhaps called a supplemental reemployment debt instrument could be used as an incentive to encourage firms to reverse the trend of replacing human capital with capital equipment in the economy. These bonds would be borrowed in the private sector and guaranteed in the public sector. They would carry the same interest tax shield as other forms of debt. The interest tax shield and the switching costs associated with repayment upon departure or dismissal of an employee would create an environment that would be more positive for reemployment of human capital. These loans could follow a life cycle format where firms in the early stages pay lower levels of debt service and in the later stages pay higher forms of debt service. It is conceivable that these could be either floating or fixed-rate notes that depend on the state of the economy.

These notes could also increase the national employment rate with only marginal implications to the economy. It is quite possible that the loss of corporate tax revenues would be offset by a much bigger gain in personal income tax associated with individuals who have been employed or reemployed by firms. Once a loan like this is established, it is conceivable that there could be a number of adjustments. Adjustments could come in the form of the status of the person before employment; were they recently unemployed? Were they part-time employed or chronically under employed? This could be used to address the case of the chronically underemployed. The size of the loans could also be adjusted to reflect the level of education. In a tiered format, each level of education could lead to larger loan amounts; one amount could be established for human capital with just a high school degree or double the scale of the loan for someone who has a college degree. It is possible that each loan value could be tripled with graduate work, experience and maintenance of professional accreditation. Thus, many social problems with regard to unemployment could be corrected with a private sector financed solution. The key question is how soon would an investor get paid back for lending money under this regime? One might consider a repackaging program that would enable these loans to be reassembled by risk grade or perhaps by level of education. Alternatively, a life cycle instrument might be employed. A life cycle instrument would allow the firm to move debt service forward depending on the economic health of the firm.

It is conceivable that a firm would hire an employee who was chronically under-employed, take the loan and establish a loan-fee payment schedule that works best for both employees and prospective employees. The switching costs associated with an employee leaving the firm should encourage firms to retain employees longer. Employed people invariably pay taxes over their working life, which will ultimately offset any loss associated with this conflict.

The reemployment bond connected directly to the lending infrastructure could be considered as a private sector initiative with government sponsored support. This stands in stark contrast to traditional proposals that typically involve tax rebates, credits or incentives. Tax rebates, credits and tax incentives drive a public-sector solution for private sector problem versus our proposal for a public sponsored private solution to a private sector problem.

As 2012 approached the US economy remained in a credit crisis – which had begun during the beginning of the recession 2007/08 lasting through this period of 5 to 6 years. During this time, the credit available to the small businessman was deeply curtailed. Banks which had been the traditional lenders to small businesses had begun to turn away from their businessmen clients in large numbers. This mandated that entrepreneurs needed to be even more creative in their search for financing. Resourcefulness and finding new solutions to their chronic funding problems became a very important item on their agenda. Alternative funding sources became increasingly popular and likely will gain even more traction over 2012. For example, long-term business owners began to turn to their own retirement programs in order to borrow funds simply to maintain their current business. In addition, many small business owners began to turn to virtual funding sites, such as Kickstarter, Profounder, IndieGoGo or Firstgiving that enable small companies to accumulate small amounts of money from investors or donors. This is prescient, since the US Congress and the Security and Exchange Commission have been discussing new laws to ease rules on such transactions (Fidelman, 2010). With these new alternative funding sources, it became increasingly clear that they would to some degree alleviate some of the problem of small businessmen obtaining financing. It is also probable that in the near term it will not be sufficient to meet the needs of many small businesses. While small businesses are collectively responsible for nearly half of all the employment in the United States, it is clear that a new way of funding has to be found. The Small Business Association (SBA) claims to channel money to small businessmen by arranging with a bank for a small loan, not to exceed approximately $250,000. However, the SBA administration does nothing more than issue a government guarantee for as much as 85% of the loan in the event of default. The problem is that this loan contains no parameters that are based on employment. If these loans were tied to employment levels of the firm, the societal benefit and tax generation benefit would be significant. During the 1970s, in response to high levels of unemployment, Congress introduced the Humphrey Hawkins Act (15 USC section 343901). The purpose of the law was to establish goals, for full employment, growth in production, price stability and balance in both trade and the federal budgets. While the act set

numerical goals to be attained by 1983 and 1988, if private enterprise could not meet the goals set forth in the act, the federal government was expressly permitted to grow in size to make a reservoir of personnel for public employment. A number of amendments to the act were added in 1979 and 1990. While the act was seriously concerned with unemployment, it did not call for explicit procedures to meet the original goals. Nonetheless, until recently significant progress was made in lowering the US unemployment rate. The act expired in 2000. It seemed that the act of measurement had led to an effect in reducing unemployment. The federal government issued a series of reports. These reports clearly showed most of the government goals had been achieved, with the possible exceptions of the downturns caused by the computer bubble of the 90s and the merger bubbles. After the act expired these successes were reversed in the downturn that began in the mid-2000s. During the current period from 2007 to present, the unemployment rate has risen by US government standards.

Since the onset of the Great Recession unemployment rates of approximately 8% to 10% have become common. The question is whether the federal government can continue to tolerate such a high degree of unemployment for such a relatively long time. It is estimated that the unemployment rate will not return to its pre-recession values until approximately 2015-16 at best. Ironically, the Humphrey Hawkins Act enabled the US government to achieve relatively modest unemployment rates. The SBA set a precedent for government underwriting of business loans without the guarantee of increasing the employment rate.

Our loan system would tie loanable funds directly to re-employment. Each loan with its corresponding guarantee would re-employ a specific individual. Ultimately, the size of the loan could be tied to level of education or specialized skills. There could be breaks of perhaps up to ten thousand dollars per degree of education, such as high school, junior college, college, graduate degree, etc. Weber (2011) has suggested that higher educational attainment leads to lower depreciation of human capital. The loan would come due upon the release of the employee from the firm. This would provide loaned funds to qualified developing and developed businesses directly connected to level of employment.

Conclusions

The goal of this paper is to address an intervention that will allow the federal government to utilize its powers without increasing the level of unemployment in the economy. We have described a system whereby the unemployment rate can be reduced significantly without any increase in the federal budget, except what may be termed a minor increase for monitoring and gathering data to make sure that the program is working appropriately.

We have discounted the possibility of direct subsidies to unemployment as a number of papers have already been written on direct subsidies to labor in order to reduce the unemployment rate. Sturgill (2003) has examples that address capital subsidies versus labor subsidies, the trade-off between capital and employment and the political economy of labor subsidies.

Capital shares and labor shares are typically treated as constant parameters in the literature. In an investigation of the role of productivity in explaining cross-country differences in output per worker, Hall and Jones (1999) assume that capital shares and labor shares remain constant across countries and equal ⅓ and ⅔ respectively. Standard measures of labor's share often entangle the fraction of income paid to human capital, a reproducible factor, and unskilled labor, a non-reproducible factor. Likewise, capital share can be split between physical capital, which is reproducible and natural capital, which is non-reproducible. The simplest labor share calculation is computed as the fraction of real GDP attributed to employee compensation. Capital's share is then computed as the residual. After all, a self-employed individual, just like an employee, is indeed a person, and the contribution to production comes from the human body. Physical capital on the other hand encompasses operating surplus, which is defined as "the excess of value added over the sum of compensation of employees, consumption of fixed capital, and net indirect taxes" by the United Nations Yearbook of National Account Statistics, and considered part of capital compensation for machines, buildings, tools, etc., and these things are inanimate, durable inputs that must be produced.

Thus, the loan program for re-employment combined with a depreciable human capital investment allowance would yield a lower structural unemployment and a well-educated, productive workforce. As we have shown, all previous programs created benefits with direct or indirect costs to the federal government. Little had been done to create a system of guaranteed loans tied to re-employment that left the government with little financial obligation (save default by target firms). Our approach would transfer the process to the private sector with guarantees by the government.

References

Arestis, Philip, Malcolm C. Sawyer. 2005, "Aggregate demand, conflict and capacity in the inflationary process," *Cambridge Journal of Economics*, Oxford University Press, volume 29, issue 6, 959-974, November.

Arrazola, María, José de Hevia, Marta Risueño, and José F. Sanz. 2005. "A Proposal to Estimate Human Capital Depreciation: Some Evidence for Spain", *Hacienda Publica Espanola*, Revista de Economia Publica 172(1), 9-22.

Becker, Gary S. 1964. *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education*, National Bureau of Economic Research.

Becker, Gary S. 1965. "A Theory of the Allocation of Time", *The Economic Journal*, Vol. 75, No. 299. (September), 493-517.

Blanchard, Olivier. 2011. *Macroeconomics Updated*, 5th edition, Prentice Hall.

Carlin, Wendy, David W. Soskice. 2006. *Macroeconomics: Imperfections, Institutions and Policies*, Oxford University Press, USA.

Chirinko, Robert S. 2008. "σ: The long and short of it", *Journal of Macroeconomics,* volume 30, 671-686.

Cobb, Charles W., Paul H. Douglas, 1928. "A Theory of Production," *American Economic Review*, volume 18 (1), March, 139-165.

Fidelman, Mark. 2010. "If This Bill Passes, The Angel Investment Community Is Dead & Companies Like Kickstarter Take Over", Business Insider, [Href: http://articles.businessinsider.com/2011-11-25/tech/30427897_1_crowdfunding-kickstarter-project-funding#ixzz1la8Ggcsm], [Last accessed 1/21/12].

Gera, Surendra. 1988. *Creating jobs in the private sector: Evidence from the Canadian employment tax credit program*, Economic Council of Canada, Minister of Supply and Services, Canada.

Hall, Robert E., Charles I. Jones. 1999. "Why do some countries produce so much more output per worker than others?" *The Quarterly Journal of Economics*, February, 83-116.

Hayek, Friedrich A. 1989. *The Collected Works of F.A. Hayek*, University of Chicago Press, USA.

Hille, Katharine. 2010. FoxConn to shift some of Apple Assembly, Financial Times, June 28 [href: http://www.ft.com/intl/cms/s/2/2429f498-82fd-11df-8b15-00144feabdc0.html#axzz1lXQGnTjd] [last accessed 1/21/12].

Foxconn to boost use of robot machines in manufacturing, BBC News, August 2, [href: http://www.bbc.co.uk/news/business-14368244] [Last accessed 1/21/12].

Kaldor, Nicholas. 1961. "Capital Accumulation and Economic Growth," *In The Theory of Capital*, edited by Friedrich A. Lutz and Douglas C. Hague, New York: St. Martin's Press.

Klump, Rainer, Olivier de La Grandville. 2000. "Economic growth and the elasticity of substitution: Two theorems and some suggestions", *American Economic Review,* volume 90, 282-291.

Organization for Economic Co-Operation and Development. 1995. Taxation, Employment, and Unemployment, Paris: OECD.

Organization for Economic Co-Operation and Development. 1996. The OECD Jobs Strategy: Pushing Ahead with the Strategy, Paris: OECD.

Rowthorn, Robert E. 1999. Unemployment, wage bargaining and capital-labor substitution, *Cambridge Journal of Economics*, Volume 23, issue 4, 413-425.

Solow, Robert M. 1980. "Employment Policy in Inflationary Times", in *Employing the Unemployed*, edited by Eli Ginzberg. Basics Books, New York.

Weber, Sylvain. 2011. "Human Capital Depreciation and Education Level: An Empirical Investigation", (November 2, 2011). Available at SSRN: http://ssrn.com/abstract=1114483 or http://dx.doi.org/10.2139/ssrn.1114483.

An Exploratory Study of Differences in Students' Views of Market 'Fairness'

John G. Marcis, 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 was 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, to many people the issue is 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 (5) 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 to have 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, Knethsch, and Thaler, 1986b, pp. 738-739).

Gorman and Kehr (1992) used 16 of the 18 scenarios developed by Kahneman et al., and created six (6) 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 (6) of the scenarios contained in Shiller et al. to 322 students enrolled in 14 sections of an "Introduction to Economics" course. Students in seven (7) sections received the survey (approximately one-half of the students) during the first week of the semester while the other seven (7) 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 (8) scenarios that were used in either the Kahneman et al. study or the Shiller et al. study. The eight (8) scenarios used in this study are presented as Table 1. Six (6) of the eight (8) 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 (2) 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 (5) 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 (8) 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 snow shovels 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 the terms "Yes" and "No" as the two possible responses. In this study, respondents were asked to respond on the "0% to 100% continuum," with "0%" indicating a scenario was "Very Unfair" and "100%" indicating the scenario was "Very Fair." It was decided that since very few issues in life related to personal perception are decided in a binary (that is, "black or white") manner, the continuum was the more robust manner in which to gather information and gauge personal 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 taking one (1) of five (5) 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.

Survey Results

A total of 181 survey instruments were used in this study, with 55 from young women and 126 from young men. 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 an area within the College of Business, 48 (27%) indicated they were planning to major in an area within the College of Liberal Arts, 39 (22%) indicated they were planning to major within the College of Fine Arts, eight (4%) indicated they were planning to major within the College of 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 six percent (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 ten percent (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 ten percent (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 ten percent (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 one percent (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 along major fields of study.

References

Carrithers, David F., and Dean Peterson. 2006. "Conflicting Views of Markets and Economic Justice: Implications for Student Learning." *Journal of Business Ethics* 69: 373-386.

Colander, David. 2003. "Integrating Sex and Drugs Into the Principles Course: Market Failures vs. Failures of Market Outcomes." *Journal of Economic Education* 34: 82-91.

Gorman, Raymond F., and James B. Kehr. 1992. "Fairness as a Constraint on Profit-Seeking: Comment." *American Economic Review* 82: 355-358.

Kahneman, Daniel, Jack Knetsch, and Richard Thaler. 1986a. "Fairness and the Assumptions of Economics." *Journal of Business* 59: S285-S300.

Kahneman, Daniel, Jack Knetsch, and Richard Thaler. 1986b. "Fairness as a Constraint on Profit-Seeking: Entitlements in the Market." *American Economic Review* 76: 728-741.

Prasad, Jyoti N., Nancy Marlow, and Richard E. Hattwick. 1998. "Gender-Based Differences in Perception of a Just Society." *Journal of Business Ethics* 17: 219-228.

Shiller, Robert J., Maxim Boycko, and Vladimir Korobov. 1991. "Popular Attitudes Toward Free Markets: The Soviet Union and the United States Compared." *American Economic Review* 81: 385-400.

Whaples, Robert. 1995. "Changes in Attitudes Among College Economics Students About the Fairness of the Market." *Journal of Economic Education* 26: 308-313.

Appendix

Table 1: "Fairness" of the Market Questionnaire

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

Very Unfair		*Unfair*		*Moderately Unfair*		*Moderately Fair*		*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	Characteristics: 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 costs of production decrease, is it fair if a business does not change the product price?	**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 by consumers?	**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 the 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***

The Long Term Effects Of NAFTA On The Economy Of The State Of Alabama

Edward T. Merkel, Troy University

Abstract

The objective of this paper is to evaluate the impact that NAFTA has had on the production of goods for export within the state of Alabama. Data listing the dollar volume of exports going from Alabama to Canada and Mexico from 1993 through 2010 are examined. Even with the decline in the total amount leaving the state since the 2008 world recession began, this study shows that state exports to NAFTA countries have grown at a rate above that of exports to the rest of the world over the last eighteen years with the largest rate of growth going to Mexico. By applying U.S. Department of Commerce computational techniques, the study estimates that since 1994 approximately 27,000 new jobs with an employment increase net of ones lost due to increased imports from Canada and Mexico of about 14,150 have been created in the state and are directly linked to rising export volumes to NAFTA nations. The paper concludes that free trade under NAFTA has enhanced and improved production and employment opportunities throughout the state, albeit at an uneven rate, during the last two decades.

Introduction

On January 1, 2012 the North American Free Trade Agreement (NAFTA) completed its nineteenth year of operation. Implemented by the governments of Canada under Conservative Party Prime Minister Mulroney, Mexico under the PRI Presidency of Carlos Salinas, and the United States under the first Democratic Clinton administration, tariffs and other duties on virtually all goods originating within and traded among the three signatory nations were phased out of existence by 1999. The NAFTA agreement also eliminated the non-tariff barrier (NTB) of quotas across the board and substantially reduced import licensing restrictions imposed by all three governments. Trade restrictions are maintained on a limited number of items deemed necessary for health and safety purposes such as pharmaceuticals when mutually agreed upon by Ottawa, Washington D.C., and Mexico City on an as needed basis. Regulations impeding transportation services, such as requiring that both trucks and drivers be registered and licensed within the country of operation, were to be abolished by 2002. However, due to the resistance of the U.S. Teamster's Union and the recalcitrance of the Clinton, Bush and Obama administrations, achieving this objective was postponed until 2011. Under a three year agreement the transportation of goods from the province of Quebec to the state of Chiapas and back on one vehicle with one driver will be allowed. And in 2007 each nation's financial institutions were given *carte blanche* to function throughout the NAFTA area as independent firms whose assets, both financial and intellectual, are protected against governmental expropriation (as predicted in Hufbauer and Schott, 1993; and confirmed in Moss, 2005 and Fleischer, et al 2011).

A free trade agreement as encompassing as NAFTA had never before been attempted among nations with such large variances in per capita incomes and overall living standards, although the expansion of the European Union into Eastern Europe did create a precedent when including many nations considered poor by German and French living standards. As reported by the World Bank (1996), when initiated in 1994 per capita GDP was $25,880 in the U.S., $19,510 in Canada, and $4,180 in Mexico. The U.S. Department of Labor (1996) further noted that average manufacturing wages in Mexico at that time were less than 10% of comparable earnings in Canada and the U.S.A. In a seminal 1991 work, Schott (1991) stressed the point that the potential life of trading blocs which include nations that have extreme differences in per capita economic conditions will be a short one due the likelihood that the majority of economic gains stemming from freer trade among member nations will flow to the poorer countries much to the chagrin and displeasure of the richer ones. These economic disparities generated heated debate which focused on the inclusion of Mexico in such an agreement. The anti-NAFTA arguments continue to this day, especially in industries and areas where production facilities and jobs seemed to have relocated to low wage Mexico, such as garment and apparel manufacturing in the U.S. Southeast (Kletzer, 2004).

Many studies both before and during the early years of NAFTA did predict that the agreement would provide a noticeable boost to the stagnating U.S. economy of the early 1990s. For example, Hufbauer and Schott (1992 and 1993), Espinoza and Noyola (1996), and Hinojosa-Ojeda (1996) ubiquitously agreed that freer trade with Canada and Mexico, already our number one and three trading partners respectively, would provide expanding export markets for U.S. producers, more jobs for

The author extends his appreciation to Ms. Olesea Voloshin at the University of Alabama for data collection assistance.

American workers in export oriented industries and related ones such as shipping, and expand consumer choice for an array of products from agricultural to labor intensive manufactured goods such as furniture. In 2001, Burfisher, Robinson, and Thierfelder (2001) confirmed this forecast when noting that U.S. exports to and imports from Canada and Mexico grew by more than 85% between 1993 and 2000, a rate greater than that with any other trading partner. Gerber (2011) lauds NAFTA as one of the principal driving forces underpinning the economic expansion in America during the 1990s as well from 2002-2007, contributing to both job growth and moderate inflation. In deference to Schott's warning, most advocates of NAFTA such as Carbaugh (2004 and 2011), Hill (2002) and Merkel and Lovik (2007) have noted that the largest share of economic benefits stemming from NAFTA have gone to Mexico. Given the size of the U.S. economy to that of Mexico, the minimal trade conducted between Canada and Mexico in tandem with the massive pre-NAFTA trade volume between the U.S.A. and Canada, and Mexico's heavy dependence on the American market for its goods, this comes as no surprise. However, these sources do agree that the benefits of NAFTA enjoyed by the U.S. economy are nonetheless positive, especially in its effects on entrepreneurship and productivity improvements. Conversely, NAFTA antagonists such as Ross Perot and Choate (1993) and the AFL-CIO (1995, 1996, 2002 and 2009) continue to cite the drain of U.S. manufacturing jobs to the Mexican economy, although of late much of this criticism has been diverted to both China and Vietnam given that average wages there are now even lower than in Mexico.

As NAFTA was being debated in Congress, the U.S. Trade Representative (1992) provided the forecast that NAFTA would increase the level of production for export and the number of export related jobs above the secular trend at that time for all fifty states, with a larger share of this growth going to those areas which enjoyed ease of shipping to Mexico and Canada. While not sharing a common border with either NAFTA nation, the port of Mobile does offer firms in Alabama a cost efficient point of shipment, especially to the port of Veracruz located on the eastern coast of Mexico. Thus, the state of Alabama has the potential for enjoying substantial economic gains from this agreement. Studies provided by the United Nations Economic Commission for Latin America and the Caribbean (1996), the Federal Reserve Bank of Dallas (Gould, 1998), and the Institute for International Economics (Hufbauer and Schott, 2004) confirm these forecasts.

The objectives of this paper are twofold. First, an evaluation is made of the impact that NAFTA has had on the production of goods for export by Alabama firms. Data listing the dollar volume of exports from Alabama to Canada and Mexico from 1993 through 2010 are examined. 1993 has been selected as the base year for analysis because it is prior to the implementation of NAFTA. Any trends in export growth can then be compared on a pre-versus post NAFTA basis so as to establish a causal link. Secondly, an estimate is provided of the net impact that NAFTA has had on the number of export generated and export related jobs based upon U.S. Department of Commerce computational techniques. The paper concludes with the observation that despite the decline in export totals to NAFTA nations after 2008, likely the result of the world-wide economic slowdown and financial crisis, overall NAFTA has boosted the economy of Alabama over the last two decades both in terms of additions to Gross State Product levels and the number of jobs generated throughout the state. Pursuant to trade barriers being removed, Alabama firms are provided with a growing source of demand for their products and Alabama labor will continue to enjoy an expanding number of jobs net of any employment losses due to increasing imports from Mexico and Canada. Additionally, should NAFTA prove to be the stepping stone to a Free Trade Area of the Americas (FTAA), state exporters would be able to reap the advantages of a free trade market with 900 million potential consumers. The recently enacted Central American Free Trade Area appears to be a positive movement in that direction.

Alabama's Exports to NAFTA: 1993 – 2010

Table 1 lists the current dollar export totals from the state of Alabama to the world and to the NAFTA nations from 1999 through 2010. These amounts are obtained from annual data provided by the U.S. Department of Commerce through the Massachusetts Institute of Social and Economic Research (MISER, 2011) at the University of Massachusetts in Amherst. Overall, exports from the state to the global economy rose from under $4 billion in 1993 to over $15.5 billion in 2010, reflecting an increase of approximately 288% from the 1993 level. Exports to the NAFTA nations increased from a bit over $1 billion in 1993 to $4.28 billion in 2010, or by about 328% of its 1993 figure. In 2008, the year prior to the current world-wide recession and financial crisis impacting trade flows, NAFTA exports peaked at $3.8 billion which was 280% of the 1993 level. However, in 2009 Alabama's total exports to the world fell by $3.5 billion and exports to NAFTA in turn decreased by $516 million. Note that by the following year 2010 both total and NAFTA exports volumes increased. In 2010 exports to the NAFTA nations rose by an amazing 981 million, just short of the one billion dollar mark, while total exports increased by about $3.15 billion. Note that exports to the NAFTA partners comprised almost one-third of this surge.

So as to remove the impact that nominal price changes may have had on export figures, table 2 provides 1993-2010 chained dollar figures, i.e., current dollar figures for 1994-2010 as listed in table 1 are deflated to reflect 1993 prices. Column 2 in table 2 lists the ratio of current year to base year (1993) prices. For example, prices in 2003 were 126.9% of prices in 1993. Thus, the current dollar total exports from Alabama in 2003 are deflated by this percentage to arrive at a chained dollar

TABLE 1: ALABAMA'S EXPORT TOTALS FROM 1993-2010

Year/To	World	NAFTA	Canada	Mexico	NAFTA Exports/ Total Exports
1993	3,979	1,034	827	207	25.99%
1994	4,654	1,230	963	267	26.43%
1995	5,407	1,378	1,138	240	25.49%
1996	5,849	1,778	1,357	421	30.39%
1997	6,702	2,476	1,463	1,013	36.94%
1998	7,037	2,231	1,608	623	31.70%
1999	6,852	2,239	1,688	551	32.68%
2000	7,918	2,506	1,758	748	31.65%
2001	7,570	2,253	1,596	657	29.76%
2002	8,267	2,350	1,688	662	28.43%
2003	8,340	2,298	1,547	751	27.55%
2004	9,063	2,561	1,831	730	28.26%
2005	10,879	3,151	2,239	918	29.01%
2006	13,899	3,229	2,269	960	23.23%
2007	14,407	3.753	2,954	799	26.05%
2008	15,879	3,815	2,785	1,030	24.03%
2009	12,355	3,299	2,572	727	26.70%
2010	15,506	4,280	3,156	1,124	27.60%

Notes: All figures are in millions of current year dollars.
Source: MISER. 2011. *State Exports by NAICS Database: Alabama*. Amherst, Massachusetts: University of Massachusetts MISER Axes Web.

TABLE 2: ALABAMA'S CHAINED DOLLAR EXPORT TOTALS FROM 1993-2010

Year	Price Index	Total Exports	Exports to NAFTA	Exports to Canada	Exports to Mexico
1993	1.000	3979	1034	827	207
1994	1.026	4536	1199	939	260
1995	1.055	5125	1306	1079	227
1996	1.086	5386	1637	1250	387
1997	1.111	6032	2229	1317	912
1998	1.128	6238	1978	1425	553
1999	1.153	5943	1942	1464	478
2000	1.192	6643	2102	1474	628
2001	1.226	6175	1838	1302	536
2002	1.245	6640	1888	1356	532
2003	1.269	6572	1811	1219	592
2004	1.299	6977	1972	1410	562
2005	1.337	8137	2361	1675	686
2006	1.391	9992	2312	1631	690
2007	1.419	10,153	2645	2082	563
2008	1.480	10,729	2577	1882	695
2009	1.481	8342	2228	1736	492
2010	1.519	10,208	2818	2078	740

Notes: All figures in millions of dollars. 1993 chained price index figures are from the author's computations.
Sources: U.S. Department of Labor, Bureau of Labor Statistics: ftp://ftp.bls.gov/pub/special.requests/cpi/cpiai.txt.on December 15, 2011; and table 1 figures.

figure of $6,572 million ($8,340 / 1.269) as listed in column 3 for the 1993-2010 time period. Similar computations are provided for chained dollar exports to NAFTA (column 4), to Canada (column 5), and to Mexico (column 6) for each year.

In chained dollar figures, up to the start of the current global economic crisis sometimes referred to as "The Great Recession", exports to Canada and Mexico combined rose from $1.034 billion to a peak of $2.577 billion in 2008, or by 149%, whereas total exports grew by a bit larger figure: 170%. This disparity was likely due to increased trade flows to China and a depreciating Canadian dollar relative to its U.S. counterpart. What is especially interesting is that between 1993 and 2008, exports to Canada increased by 127% while exports to Mexico increased by a rate almost twice as large, or by 236%. Trade flows to both the world and the NAFTA partners fell precipitously in 2009 which comes as no surprise given the abysmal state of the world economy that year. However, exports to both Canada and Mexico resumed an upward trend in 2010; in fact by the end of that year exports to Canada and Mexico were at an all time high of over $2.8 billion, 172% above the 1993 figure. While chained dollar sales to Canada increased by 151%, those to Mexico rose by an incredible 257%. Undoubtedly the quick resurgence of NAFTA trade helped to underpin the fact that Alabama's overall unemployment rate was more than a full percentage lower than the national figure during 2011. All indications are that export activity to our NAFTA partners will rise even further in 2011. Whether current (table 1) or chained (table 2) dollar figures are evaluated, and notwithstanding declining exports to Mexico and to Canada during the 2007-9 period, the following observations can be made: NAFTA has become a dominant market for Alabama's export industries and the state's exports to Mexico have increased much more rapidly than those to Canada. Apparently, the adverse impact that NAFTA was predicted to have on Alabama's economy has simply not materialized.

The Impact of NAFTA on Alabama Employment

Opponents to NAFTA are quick to point out the deleterious effect that free trade, especially with low wage nations such as Mexico, India, and China, will have on the U.S. job market. Perot and Choate (1993) and the AFL-CIO (1995) predicted the potential loss of 150,000 to 250,000 jobs primarily impacting low to semi-skilled industrial workers and farm labor. Since Alabama is a major producer of textiles and apparel, it was feared that free trade with Mexico under NAFTA would decimate these industries within a few years (Kletzer, 2004). That great sucking sound made by American jobs being lost to the low wage market of Mexico *a la* Ross Perot's infamous comment was believed to be imminent. However, employment figures for Alabama pursuant to NAFTA show an entirely different net job market effect.

Referring to the chained dollar figures provided in table 2 above, Alabama's exports to NAFTA countries have increased by about $1.784 billion from 1993 through 2010. Estimates of the total number of new jobs that each additional $1 billion of exports will create in the U.S. economy range from the U.S. Trade Representative (1992) figure of 10,000 to the Hufbauer and Schott (1992) number of 14,500 to the 20,000 job increase forecast provided by Hinojosa-Ojeda (1996). Using the median figure of 15,000 jobs created for each additional $1 billion in exports, the overall $1.784 billion chained dollar increase in goods sold to NAFTA members since 1994 has likely generated about 27,000 new jobs in the state. This figure would include employment created directly in expanding export industries such as electronics and transportation equipment and jobs created in those service industries which would complement rising exports such as shipping, insurance, and finance.

Without question, NAFTA has also likely produced some job losses in the state within those industries competing with expanding imports. A prime example held up by NAFTA critics is the decline in the state's textile and apparel employment figures. In 1998, for example, the Center for Business and Economic Research at the University of Alabama (1998) estimated that between 1994 and 1997, 13,000 apparel manufacturing jobs were lost in Alabama due to the closing of a number of "cut-and-sew" plants. The Department of Industrial Relations for the State of Alabama (2002) puts the job loss figure in this industry at about 12,500 between 1995 and 2000.

In allowing for jobs lost due to rising imports from NAFTA trade, table 3 contains data on the impact that trade has had upon *net* job totals throughout the state between 1994 and 2010. Using the U.S. Department of Commerce estimate that 1 of every 8 new jobs in the U.S. labor market can be linked to export activity (cited in Jones, 1998 and Carbaugh, 1998), the total number of jobs created in the state during each year attributable to total exports net of any job losses due to import competition are estimated in row three. By multiplying these figures by the percentage of exports sent to NAFTA countries, annual estimates of the jobs created in the state related to trading activity with Mexico and Canada are found in column five. It appears that despite employment losses in industries competing with products from NAFTA nations, especially those from Mexico such as apparel, textiles, and furniture, Alabama's labor market gained over 14,000 new jobs from this trade agreement during its first ten years of operation. Other studies ascertaining the net job effect of free trade under NAFTA on southeastern states, such as Jones (1998) and Merkel (2000), and on states in the far western region (Hinojosa-Ojeda, 1996), also show that employment gains from NAFTA have consistently exceeded job losses.

TABLE 3: ALABAMA'S EXPORT SUPPORTED JOBS ATTRIBUTABLE TO NAFTA 1994-2010

Year	Total Job Increase x 1/8=	Export Jobs	x NAFTA/Total Exports	= NAFTA Jobs
1994	42	5.25	26.43%	1.39
1995	45	5.63	25.49%	1.44
1996	25	3.13	30.39%	0.95
1997	20	2.50	36.94%	0.92
1998	51	6.38	31.70%	2.02
1999	45	5.63	32.68%	1.84
2000	18	2.25	31.65%	0.71
2001	(-11)	(-1.38)	29.76%	(-0.41)
2002	6	0.75	28.43%	0.21
2003	1	0.13	27.55%	0.04
2004	25	3.13	28.26%	0.88
2005	62	7.75	29.01%	2.25
2006	35	4.38	23.23%	1.02
2007	53	6.63	26.05%	1.73
2008	6	0.75	24.03%	0.18
2009	(-20)	(-2.50)	26.78%	(-0.67)
2010	(-10)	(-1.25)	27.60%	(-0.35)
			TOTAL	14.15

Notes: All employment figures rounded in 1000's and are in the non-agricultural sector of the state.

Sources: Department of Industrial Relations. Various Issues. *Alabama Labor Market News*. Montgomery, Alabama: State of Alabama; and table 1 figures.

With regards to the negative impact that NAFTA may have had on apparel jobs throughout the state, always a sore point advertised by NAFTA opponents, note that employment losses within this industry merely continue a long term secular decline, the origin of which preceded NAFTA by decades. Hinojosa-Ojeda (1996) reports that the number of jobs in the U.S. apparel industry began to decline in 1979, 15 years *prior* to the NAFTA agreement. By 1995, the industry was employing 27% fewer persons than it was in 1979. Kletzer (2005) notes emphatically that the apparel manufacturing job losses were primarily due to the application and implementation of mechanical and technological advances in the U.S. economy. In short, machines and in many cases robots replaced human labor in the assembling of clothing. Lindsey (2004, p.1) reaffirms this observation when stating: "There is no significant difference between jobs lost because of trade and those lost because of new technologies. All those job losses are a painful but necessary part of the larger process of innovation and productivity increases that is the source of new wealth and rising living standards".

Therefore, rising imports have been a minor causative element of any job losses throughout this industry during the NAFTA period. The fact that jobs were lost prior to any impact that NAFTA could have had on the nation's, and Alabama's, apparel manufacturers implies that no, or at most a minimal, connection between free trade with Mexico and Canada and declining apparel employment levels can be inferred. Hinojosa-Ojeda (1996, p. 25) further discredits the argument that free trade with Mexico since 1994 has shrunk the U.S. apparel making job market when stating: "The vast majority of recent job losses [in apparel making] are due to a fall in export [growth] with a much smaller amount due to the increase in competitive imports". The fall in exports to Mexico and Canada at the start of the last decade and in 2009 noted in table 1 above were primarily due to an appreciating dollar and the 2001 and 2008 recessions. These events in tandem with the industry becoming more capital intensive in nature, thus not the much feared and touted surge in clothing imports, could well explain most of the recent job losses in Alabama's apparel industry.

Conclusion

The overall impact of NAFTA on the economy of the state of Alabama appears to have been positive. Exports to Canada and Mexico have grown faster than those to the rest of the world since 1994 with the highest export growth being with Mexico. In fact, in the year 2000 Mexico became Alabama's number two trading partner by replacing Japan and has vacillated between the number two and three slots ever since. Trade outflows to NAFTA countries contributed to a bit more than 1% of Gross State Product (GSP) in 1993. In 2010 that percentage doubled to approximately 2% (Bureau of Economic

Analysis, various issues). While some Alabama jobs may have been lost due to competitive imports originating in the NAFTA nations, 27,000 new jobs with an estimated net employment increase of over 14,150 jobs can be attributed to freer trade with Canada and Mexico. These statistics belie the on-going criticism that free trade in general, and NAFTA in particular, has received in recent years. The 2008 recession has renewed the cry for protectionism from increased agricultural subsidies to greater restrictions placed on imported lumber, wearing apparel, and steel. This is one of the major reasons why the current Doha round of World Trade Organization negotiations collapsed in Cancun, Mexico in 2004, in Geneva in 2006, and again in Cancun in 2008, and why the discussion on establishing a Free Trade Area of the Americas (FTAA) is currently at an impasse. However, as Lindsey (2004, p. 1) points out: "Calls for new trade restrictions to preserve current jobs are misguided".

In conclusion, all residents of Alabama need to understand and appreciate the economic advantages that could be gained through free trade under NAFTA types of agreements. The Center for International Trade and Commerce (1998, p. 1) in Mobile, Alabama emphasizes this warning with the following statement: "[As] a result of increased international activity the South is now the fastest growing region in the United States in the area of import/export and reverse investment". While this statement was made during the crest of the 1990s business cycle, the 2008 recession adversely impacted the South in general and Alabama in particular much less than it did in the Northeast and Midwest regions of the country. In fact, unemployment in the state remained noticeably below the national rate until February, 2011 (*Economic Update,* July-August 2011). Much of this mitigated effect was due to international trade flows, in particular to its NAFTA customers.

Ghemawat (2011), in a sweeping analysis of the positive impact that trade agreements have had on member nation's economies since the early 1950s, argues that two nations will engage in 47% more trade if both belong to a trading block such as NAFTA than if they do not (also see Bagwell and Staiger, 2011). Of particular interest is that trade increases by 114% if a common currency is introduced. Maggi (1999) stresses that a trading bloc is a group of countries that value longer term economic cooperation among themselves more than any short term transient gains that could be realized via the removal of barriers to trade flows. Therefore, if a successful NAFTA does become the stepping stone towards a FTAA by lessening the current resistance to the latter, possibly supplemented with the ubiquitous adoption of a single form of money throughout the Americas, then Alabama producers could be provided with an expanded market of almost 900 million consumers throughout all of North, Central, and South America and the Caribbean basin (World Bank, 2004). Pursuant to the positive impact which NAFTA has provided to Alabama's economy, an expanded FTAA could only lead to further improvements in employment opportunities and living standards throughout the entire state. This prediction is noted in a 2000 study evaluating the economic impact of the first five years of NAFTA (Merkel, 2000, p. 20): "So as to fully reap the advantages of free trade, the citizenry of the state of Alabama must therefore embrace, and not fear nor resist, open borders throughout the Americas."

References

AFL-CIO. 1995. *NAFTA Math – 21 Months Later.* Washington, D.C.: AFL-CIO.

_____. 1996. *NAFTA Math – 2 Years Later.* Washington, D.C.: AFL-CIO.

_____. 2002. *Employment Losses Due to Unfair Trade.* Washington, D.C.: AFL-CIO.

_____. 2009. *Nafta and the Great Recession of 2008.* Washington, D.C.: AFL-CIO.

Bagwell, Kyle & Robert Staiger. "What Do Trade Negotiators Negotiate About? Evidence from the World Trade Organization". *American Economic Review* 101, No. 4: 1238-1273.

Bureau of Economic Analysis. Various issues. *Survey of Current Business.* Washington, D.C.: U.S. Government Printing Office.

Burfisher, Mary, Sherman Robinson, and Karen Thierfelder. 2001, Winter. "The Impact of NAFTA on the United States". *Journal of Economic Perspectives* 15, No.1: 125-144.

Carbaugh, Robert. 2011, 2004 & 1998. *International Economics.* Cincinnati, Ohio: Southwestern Publishing Company.

Center for Business and Economic Research. Various issues. *Alabama Business and Economic Indicators.* Tuscaloosa, Alabama: University of Alabama.

Center for International Trade and Commerce. 1998, September. *World Trade Update.* Mobile, Alabama: University of South Alabama.

Department of Industrial Relations. Various issues. *Alabama Labor Market News.* Montgomery, Alabama: State of Alabama.

Economic Update. July-August 2011. Jacksonville State University. 21, No. 4: 5.

Espinoza, Enrique and Pedro Noyola. 1996. *Emerging Patters of Mexican – U.S. Trade.* Washington, D.C.: Brookings Institute.

Fleischer, Petra et al. 2011. "A Bayesian Analysis of Market Information Linkage among NAFTA countries using a Multivariate Stochastic Model. *Journal of Economics and Finance*. 35, No. 2, April: 123-148.

Gerber, James. 2011. *International Economics.* Boston: Addison-Wesley.

Ghemawat, Pankaj. 2011. *World 3.0: Global Prosperity and How to Achieve It.* Boston: Harvard University Press.

Gould, David. 1998. "Has NAFTA Changed North American Trade?" *Economic Review-Federal Reserve Bank of Dallas* 1st Quarter: 12-21.

Hill, Charles. 2002. *Global Business Today.* Boston: Irwin-McGraw Hill.

Hinojosa-Ojeda, Raul. 1996. *North-American Integration Three Years After NAFTA.* Los Angeles: UCLA North American Integration and development Center.

Hufbauer, Gary and Jeffrey Schott. 1992. *North American Free Trade: Issues and Recommendations.* Washington, D.C.: Institute for International Economics.

_____. 1993. *NAFTA: An Assessment.* Washington, D.C.: Institute for International Economics.

_____. 2004. *NAFTA: A Ten Year Appraisal.* Washington, D.C.: Institute for International Economics.

_____. 2010. *NAFTA: After Fifteen Years.* Washington, D.C.: Institute for International Economics.

Jones, Marcia. 1998, Winter. "The Impact of NAFTA on Jobs in Georgia". *Southern Economic Developer* 8, No. 1: 7-10.

Kletzer, Lori. 2004. *Workers at Risk: Job Loss from Apparel, Textiles, Footwear, and Furniture.* Washington, D.C.: Institute for International Economics.

Lindsey, Brink. 2004, March 17. "Job Losses and Trade". *Cato Institute Trade Briefing Paper* 19: 1-12.

Maggi, Giovanni. 1999, March. "The Role of Multilateral Institutions in International Trade Cooperation". *American Economic Review* 89, No. 1: 190-214.

Merkel, Edward. 2000, Winter. "The Impact of NAFTA on the Economy of the State of Alabama: The First Five Years Bring Good News". *Business and Economic Review* 24, No. 1: 15-20.

_____ & Larry Lovik. 2007. "Mexico within NAFTA: The Silent Winner". *Academy of Economics and Finance Papers and Proceedings*: 256-261.

Massachusetts Institute of Economic and Social Research. 2011. *State Exports by NAICS Database: Alabama.* Amherst, Massachusetts: University of Massachusetts.

Moss, Joanna. 2005. "Economic Developments During NAFTA"s First Decade" in *International Economics and International Economic Policy,* ed. Phillip King and Sharmila King. New York: The McGraw-Hill Companies.

Perot, Ross and Pat Choate. 1993. *Save Your Job, Save Your Country: Why NAFTA Must be Stopped Now.* New York: Hyperion, 1993.

Schott, Jeffrey. 1991, March. "Trading Blocks and the World Trading System". *The World Economy* 14, No. 1: 1-17.

United Nations Economic Commission for Latin America and the Caribbean (UNECLAC). 1996. *NAFTA Implementation in The U.S.: First Two Years.* Washington, D.C.: UNECLAC.

U.S. Department of Labor, Bureau of Labor Statistics. 1996. *International Comparison of Hourly Compensation Costs for Production Workers in Manufacturing.* Washington, D.C.: U.S. Government Printing Office.

_____. 2011, December 15. ftp://ftp.bls.gov/pub/specialrequests/cpi/cpiai.txt.

U.S. Trade Representative. 1992. *NAFTA will Expand U.S. Exports and Generate Jobs.* Washington, D.C.: U.S. Trade Representative Office.

World Bank. Various Issues. *World Development Report.* New York: Oxford University Press.

Economic Impact of Free Trade on Mexico

Pat O'Brien, University of Southern Mississippi

Abstract

Addressing motivations for population flows, this research compares the Mexican economy prior to the commencement of a free trade regime under the North American Free Trade Agreement (NAFTA), in January of 1994, with changes in the Mexican economy as it phased in implementation of that treaty. Presented in a consolidated format, this paper addresses how and to what extent this free trade regime affected standards of living of the average Mexican citizen and his/her motivation to seek better locations of opportunity.

Introduction

This is an analysis of the impact of increasing free trade since 1980 on the welfare of the citizens of Mexico. Are the citizens of Mexico economically better off with the growing Mexican regime of free trade and the implementation of the North American Free Trade Agreement (NAFTA) than they would have been without it?

Allee, *et al.* (1949, p. 539) point out one has three choices when exposed to adversity, "die, adjust, or migrate." Ravenstein (1885) demonstrates the centrality of financial incentives in the migration. Did free trade contribute significantly to financial and other incentives that may have been instrumental in emigration from Mexico? In addition to emigration rates as a proxy for population welfare, individual and personal forms of measurement advocated by the 2009 Report by the Commission on Measurement of Economic Performance and Social Progress (Stiglitz, *et al.*, 2009) are used to measure personal welfare. Mexican international trade is compared against these measures of population welfare? Data are assessed from prior to the implementation of NAFTA with its original signing, December 8, 1993, though the present. Recent years are included, in spite of the volatility associated with the most recent worldwide recession. With few noted exceptions, data are drawn from the World Bank and UNESCO databases. Regression analysis of indicators of well-being is used to confirm crucial trends in personal well-being of the average Mexican citizen with increased free trade.

Literature review

Concerns about NAFTA, and the subsequent impacts of freer trade, have varied throughout its tenure. Krugman (1993) argues that NAFTA is a foreign policy issue with no macroeconomic affect. Nápoles (2007) posits that failures of earlier Mexican economic policies resulted in the Mexican recession of 1982, heavy foreign debt, and a foreign exchange crisis that forced the government to adopt liberal trade policies in hope of economic recovery and contends that unrestricted trade does not benefit Mexico. Diep (2008) notes that the optimism of the national leadership of the member nations toward NAFTA is not shared by constituents and that NAFTA's survival is primarily a political question.

Perspectives also vary about the design, implementation, and effectiveness of NAFTA to meet the needs of the citizens of its member nations. Bhagwati (1994, pp. 102-103) addresses how the imperfections of factor and product markets provide the intellectual support to protectionist forces. Bhagwati reminds us of Britain's public disbelief in free trade following the unemployment of 1930's, and that John Maynard Keynes promoted the use of optimal tariffs. Fayyaz (2008) notes that peasant perception of a widened the rich-poor gap in Mexico. Beaulieu (2000) documents negative effects on employee income following free trade implementation. Ederington (2007, p. 239) cautions that trade liberalization will attract "polluting industries from high income/stringent environmental regulation countries to low income/lax environmental regulation countries." White (2007) describes United States policies of protectionism and security that limited its participation in NAFTA. Shaw (2008) supports White's assertion by describing legal battles related to antidumping and countervailing duties under NAFTA. Piérola and Horlick (2007) suggest that implementation of regional trade agreements must be sensitive to the needs, rights of the citizens of participating countries. Carlsen (2008) contends that NAFTA is locked into economic neoliberalism. She accepts the positive aspects of an open market and an export-oriented economy, but identifies a downward pressure on wages, use of natural resources, and suppression of dissent in Mexico.

There are also forces at work within and following the implementation of free trade that have positive impact on citizen welfare and others that have no perceived effect. Bhagwati (2004) contends that it is the government's responsibility to minimize adverse effects on citizens. Barbier and Hultberg (2007) feel that decisions to relocate operations as countries integrate economically are affected by "the relative marginal cost of production between the regions, the cost of relocation, [and] the cost of exporting its goods across borders." This occurs within as well as between countries. Delgado-Wise and Covarrubias (2007) describe the restructuring of the Mexican labor force with the implementation of free trade. Hufbauer and

Schott (2008) contend that the trade regime under NAFTA is successful; tripling trade. They note that there are winners and losers. They, like Bhagwati, site governmental failure to deliver tax and energy reforms, and failure to fund education infrastructure as responsible for comparatively slower economic growth in Mexico. They urge a change in the discussion focus of NAFTA from political rhetoric to "thoughtful dialogue," and recommend updating and improving the treaty to serve the interests of the member nations in the areas of labor, environment, climate change, and border security.

To better understand the impact of economic policy since, "what we measure shapes what we strive to pursue – and what we pursue determines what we measure," (Stiglitz, 2009, p. 9) some of the issues about free trade may relate to what and how we measure success. The Report by the Commission on Measurement of Economic Performance and Social Progress (Stiglitz, *et al.*, 2009) commissioned in February, 2008, by the President of the French Republic, Nicholas Sarkozy and chaired by Joseph E. Stiglitz, with chair advisor Amartya Sen provides discerning approach to how and what we measure in an effort to better assess the well-being of a population. It provides twelve recommendations for change in "evaluating material well-being." This report is fundamentally about changing forms of measurement. The report recommends assessing income and consumption rather than production, focusing on the household rather than the nation or region, advocates median rather than average measures, and urges the use of subjective measures of well-being as well as monetary measures. The report recommends the use of GNI per capita rather than GDP, as it accounts for depreciation of assets and income from abroad, and more effectively measures net wealth. The Gini coefficient also better reflects population income, than national averages, as it addresses distribution. Household income and consumption indicators better reflect population condition than average national statistics. Health and education measures also indicate changes in population well-being. The report recommends a fundamental change is the types of data collected for use in public policy decisions. It addresses eight key dimensions of well-being; "material living standards (income, consumption and wealth) . . . health . . . education . . . personal activities including work . . . political voice and governance . . . social connections and relationships . . . environment (present and future conditions) . . . insecurity, of an economic as well as a physical nature" (Stiglitz, *et al.*, 2009, p. 13-15). Using these recommendations is a guide a broad approach to understanding the true effect of freer trade on the citizens of Mexico.

Theoretical framework

This research addresses Mexican population well-being from the perspective of dimensions of well-being and treats the Mexican population and economy as a whole, relating to the impact of freer trade. Todaro (2006) notes that "gross national income (GNI) per capita, the most commonly used measure of the overall level of economic activity, is often used as a summary index of the relative economic well-being of the people in different nations. Stiglitz (2009) argues that national economic statistics "originally intended as a measure of market economic activity, including the public sector, have increasingly [and wrongly] been thought of as measure of societal well-being."

Changes in free trade; Mexican exports, total Mexican international trade, and net Mexican emigration is expected to effect material living standards, household income and consumption, health, and education; the well-being of the population. Taken together these results help provide insight into the well-being of the Mexican population and subsequent compounding incentives to emigrate.

The independent variable for exports drawn from the World Bank Database (2012) is "exports of goods and services (current US$)." The independent variable for total trade is derived from the sum of "imports of goods and services (current US$)" and the "exports of goods and services (current US$)." In all cases, net Mexican migration is supplied as an independent variable to measure the depth of individual relation to the dependent measures of well-being. United States Census Bureau data is used to supplement World Bank data since net migration is only available on quinquennially from the World Bank. Regressions are conducted individually on dependent variables.

Regression analysis of exports, total trade, and net Mexican migration is used to predict changes to GNI per capita PPP, which is used as a proxy for material living standards.

$$Y' = a + b_1X_1 + b_2X_2 + b_3X_3 \quad (1)$$

where Y' is GNI per capita PPP, "a" the constant, b_1, b_2, and b_3 are regression coefficients, and X_1, X_2, and X_3 are exports, total trade and net migration respectively. GNI per capita PPP is from the World Bank Database (2012). Though this measure is an average and so does not provide a median income measure as suggested by Stiglitz, it does provide a general index of well-being comparable to citizens of other nations.

Regression analysis of exports, total exports, and net Mexican migration is used to predict changes in household income and consumption.

$$Y' = a + b_1X_1 + b_2X_2 + b_3X_3 \quad (2)$$

where Y' are individual measures of household income and consumption, "a" the constant, b_1, b_2, and b_3 are regression coefficients, and X_1, X_2, and X_3 are exports, total trade and net migration respectively. A comprehensive measure of household income and consumption is "household final consumption expenditure (constant 2000 US$)," from the World Bank Database (2012). Labor participation rate and annual consumer price inflation rate are separately treated as dependent variables in this analysis. Note that consistent with Stiglitz, *et al.* (2009) recommendation GDP per capita and other summative macroeconomic data are ignored.

Regression analysis is conducted using exports, total trade, and net Mexican migration to predict changes to health.

$$Y' = a + b_1X_1 + b_2X_2 + b_3X_3 \quad (3)$$

where Y' are individual measures of health, "a" the constant, b_1, b_2, and b_3 are regression coefficients, and X_1, X_2, and X_3 are exports, total trade and net migration respectively. The most comprehensive measure of health is "life expectancy at birth, total (years)," is drawn from the World Bank Database (2012) and is used as the dependent variable in this analysis. Other dependent variables examined include infant mortality rate per 1,000 live births, mortality under 5 years per 1,000 live births, percent of DPT immunization for children ages 12-23 months, percent of measles immunization of children ages 12-23 months, and fertility rate. Other data related to public health is not consistently and reliably available for each year since 1980 and so is not useful for this analysis.

Regression analysis is conducted using exports, total trade, and net Mexican migration to predict changes to education.

$$Y' = a + b_1X_1 + b_2X_2 + b_3X_3 \quad (4)$$

where Y' are individual measures of educational attainment, "a" the constant, b_1, b_2, and b_3 are regression coefficients, and X_1, X_2, and X_3 are exports, total trade and net migration respectively. The most comprehensive measure of educational attainment is "literacy rate, adult total (% of people ages 15 and above). It is important to note that although adult literacy is a strong indicator of the successful deployment of a public education system, these figures are only available in the decennial years and so are not useful in this analysis. Other dependent variables used as proxies for educational attainment include, primary completion rate; primary, secondary and tertiary school enrollment; and female to male ratios in primary and secondary schools.

Research question

"Is the average citizen of Mexico is better off under an increasingly free trade regime?" To the extent that the people of Mexico are not better off, the incentives to migrate increase. It is expected from free trade models noted by Sawyer, *et al.* (2004) and Trefler (2004) that measures of economic growth increase as free trade increases. These measures of well-being include financial as well as subjective and general welfare indicators. The dependent variables are proxies for economic and personal well-being. GNI per capita PPP is used to measure economic well-being. Household final consumption expenditures, labor participation rate, inflation of consumer price levels, and unemployment rate measure household income and consumption. Life expectancy, infant mortality rate, under 5 mortality rate, DPT and Measles immunization rates and fertility rate are used as proxies to measure health. School enrollment, completion rates and female-to-male attendance ratios are used as proxies to measure education efficacy. Regression analysis measuring the extent of population well-being is tracked from 1980 to present. The null hypothesis posits that as free trade increased, no significant improvements or loss of citizen well-being occurred. Freer trade under NAFTA is assumed to result in increases in population well-being and reduced emigration incentives. It is also posited (null hypothesis) that increases in net emigration minimally affect population well-being. The research hypotheses are that as free trade (or net emigration) increased, improvements of citizen well-being (H_1) occurred, and that as free trade (or net emigration) increased, the loss of citizen well-being (H_2) occurred.

H_0: Increases in international trade (and net emigration) do not or only moderately affect measures of population well-being.

H_1: Increases in international trade (or net emigration) positively affect measures of population well-being.

H_2: Increases in international trade (or net emigration) negatively affect measures of population well-being.

Exports and total trade are used as independent variables to measure the extent of free trade. Net Mexican migration is also treated as an independent variable to evaluate its relationship with measures of population well-being (dependent variables).

Findings

The World Bank data Statistics of Mexico, (2012) provides development indicators for that country including economic and quality of life assessments. Sources of this data are problematic in that World Bank Statistics tends to provide data with unreported years for many of the data elements. Other data sources such as the United States Census Bureau, UNESCO and Mundi.com are used to complete these data sets. A few variables lacked a single year and these were extrapolated from the existing data set. Other issues include that some data is gathered primarily for market or production purposes, (Stiglitz, 2009) and often the data is not gathered at all.

Table: Regression Results

Summary	R Square	Adjusted R Square	Model Significance	Intercept	*Export Coefficients*	*Export P-value*	*Total Trade Coefficients*	*Total Trade P-value*	*Net Migration Coefficients*	*Net Migration P-value*
GNI per capita, PPP	0.975	0.973	0.000	4202.368	-0.035	0.200	0.034	0.014	1321.635	0.233
Household final consumption expenditure (constant 2000 US$)	0.984	0.982	0.000	185,946,000,000	(1,553,909)	0.026	1270118	0.000	(32,349,377,223)	0.252
Labor participation rate %	0.788	0.764	0.000	56.554	0.000	0.534	0.000	0.357	-3.657	0.034
Annual consumer prices inflation %,	0.671	0.635	0.000	73.496	0.004	0.001	-0.002	0.001	71.963	0.093
Life Expectancy (years)	0.923	0.914	0.000	66.049	0.000	0.888	0.000	0.528	-9.382	0.000
Infant mortality /1,000 live births	0.926	0.917	0.000	57.271	0.000	0.909	0.000	0.684	36.377	0.000
Under 5 Mortality /1,000 live births	0.922	0.914	0.000	75.414	0.000	0.980	0.000	0.638	52.165	0.000
DPT Immunization %	0.788	0.765	0.000	31.289	-0.001	0.161	0.000	0.115	-93.502	0.000
Measles Immunization %	0.778	0.754	0.000	22.473	-0.001	0.112	0.001	0.089	-123.155	0.000
Fertillty rate (births per woman)	0.867	0.852	0.000	4.718	0.000	0.883	0.000	0.876	2.841	0.000
Primary School completion rate %	0.871	0.856	0.000	83.004	0.000	0.006	0.000	0.018	-14.337	0.005
Primary School Enrollment Rate	0.880	0.866	0.000	122.524	0.000	0.037	0.000	0.032	21.042	0.000
Secondary School Enrollment Rate	0.947	0.941	0.000	44.520	0.001	0.002	0.000	0.013	-12.989	0.055
Tertiary School Enrollment Rate	0.937	0.929	0.000	11.930	0.000	0.003	0.000	0.025	-2.558	0.362
Ratio female to male, primary enrollment	0.481	0.421	0.001	96.810	0.000	0.619	0.000	0.489	-0.079	0.952
Ratio female to male, secondary enrollment	0.832	0.812	0.000	85.781	0.000	0.860	0.000	0.629	-20.547	0.001

Regression of export value, total export value and net Mexican migration as causal variables are assessed against sixteen dependent variables that represent the domains of well-being. The "ratio of female to male; primary enrollment," lacks a significant R-square value and since this indicates that a substantial portion of the variance within the model cannot be explained this variable is ignored.

Four gross economic, household and consumption indicators are assessed. GNI per capita PPP shows explanation of model variance and significance, but only the independent variable "total trade" demonstrates a significant statistical probability (p-value) that the coefficient measures the relationship between the two variables. If the "total trade" increases by $1 million, then GNI per capita PPP will increase by 3.4 cents. Household final consumption expenditure, labor participation rate, and consumer price inflation explain variance and demonstrate model significance as well. Acceptable p-values for "exports" and "total trade" exist for household final consumption expenditure and consumer price inflation. If exports increase by $1 million, then total Mexican household final consumption will decrease by $1,553,908 and consumer price inflation will increase by .4 percent. A $1 million increase in total trade will result in a $1,270,117 increase in total Mexican household final consumption and a .2 percent decrease in consumer price inflation. An annual increase of 1 million emigrants historically decreases labor participation rate by 3.6 percent.

Six indicators of health are assessed. All of the health variables explain model variance and significance. P-values, however, demonstrate reasonable statistical probability only for "net migration." An annual increase of 1 million emigrants predict a decrease of life expectancy of 9.3 years, an increase in infant mortality of 36/1,000 live births, an increase of under 5 mortality of 52/1,000 live births, a significant decrease in immunizations; 93% and 123%, and a fertility increase of 2.8 children.

Indicators of education are assessed. All of these indicators except the ratio of female to male, primary enrollment explain model variance and significance. Although primary school completion rate, primary, secondary and tertiary school enrollment have acceptable p-values for trade, the related coefficients are so small that changes in "exports" or "total trade have no effect on these dependent variables. Net migration data shows that acceptable p-values allow that an annual increase

of 1 million emigrants annually will result in a 14 percent drop in primary school completion rate, a 21 percent increase in primary school enrollment, and a 20 percent drop in the female to male ration of secondary enrollment.

Analysis

The growth in Mexican Gross Domestic Product (GDP) has been constant; approximately 7% annually prior to NAFTA and 12% afterward. International trade, as a percent of GDP doubled following the implementation of NAFTA. Increases in exports have been predicted to increase the wealth and well-being of the average Mexican citizen, but as economies grow, industrial sectors change and jobs and social security suffer. Imports grew at a faster rate than exports, but exports and total trade continued to climb throughout the period of study. Variation in trade prior to NAFTA has been replaced with consistent growth. Rates of emigration over the period studied varies randomly, but roughly averages just below 400,000 annually.

An assessment of these measures of well-being from 1980 included data from years since to onset of the recent worldwide recession and impacts the relation among the variables studied. The limited effect of exports and total trade on GNI is consistent Bulmer-Thomas' (2003, p. 50) view of the unrealistic burden of export led growth as a sole source of a nation's economic development. We cannot reject the null hypotheses for this measure as there were no significant increases or losses in population well-being as trade grew.

Three measures of household income and consumption also provide inconclusive results. Trade directly affects household final consumption expenditures as exports increase household final consumption expenditure decreases and inflation marginally rises. As total trade increases household final consumption expenditure increase and inflation marginally drops.

The six measures of health are unaffected by exports or total trade, but migration affects all of them. As net migration increases life expectancy and DPT immunization rates significantly decrease, and infant mortality and under 5 mortality increase significantly. It is important to note that in this study migration is measured at 1 million, historically this figure is closer to 400,000 persons.

Six measures of education, as components of well-being, though significant, affect little change. As noted, the ratio of female to male, primary enrollment is ignored. Although related to exports and total trade, though enrollment rates and primary completion rates are statistically significant, the coefficients are so small that there is minimal affect. Increases in emigration negatively affect primary completion rates while increasing enrollment and decreasing the female to male ratio in secondary schools.

In summary, although we cannot reject the null hypothesis for the impact of international trade, there are consistent variables that show a negative impact on population welfare as net emigration increases. The null hypothesis is rejected in favor of H_2. There may be several causes of changes in personal welfare in Mexico, but net emigration may affect measures of well-being deeper and in more ways than international economic exchange. Lack of sufficient, reliable data remains a challenge. Many of these measures did not possess statistical significance within the model. Other measures of economic and societal condition from the endogenous, international and academic sources are excluded because the data collected does not effectively span the timeframe of this study or are missing data on a random or consistent basis.

Conclusions

Simon Kuznets's (1955) curve describing the rise and fall of inequality in the path of increased a nation's economic growth predict changes to the population of Mexico as its economy grows. Population devoted to agriculture is being replaced by manufacturing and it by services. As these changes occur inequality is expected. These changes in major industries have an adverse impact on the population contributing to emigration incentives. To the extent that any general increase in well-being in Mexico has come, it has at a cost to some industry sectors. Although services and merchandise sales sectors in Mexico have increased, manufacturing continues to decline. This change in the complexion of the Mexican economy is consistent with those of developed countries as labor shifted from the agricultural sector to manufacturing and subsequently to services, displacing some and given opportunity to others. Restructuring of the labor force is in progress. As Allee and Ravenstein predict, these pressures account for population movement within and from Mexico to locations of opportunity; to larger metropolitan areas within Mexico and to the United States.

International trade and Gross Domestic Product (GDP) have increased since implementing the NAFTA, but only at paces moderately faster than prior to the treaty. The consumption of utilities and imports has grown. Per capita gross domestic income (GNI) measured in purchasing power parity (PPP) has grown and the Gini coefficient, a measure of income inequality, has dropped (~50% to ~46%). Mexican citizens are, on average, (a keyword) economically better off after the implementation of NAFTA. Using Ravenstein's (1885) assessment that financial incentives are primary motivations for the

migration of people groups, NAFTA's free trade regime cannot be said to have significantly contributed to the general emigration from Mexico.

Recommendations for changes in Mexican public trade policy and industry growth must be cautioned by these findings. The free trade regime may have been good for the Mexican population, but infrastructure development implemented as part of the economic development policy must include infrastructure improvement, incentives to export manufacturing products, improving technology, eliminating corruption within the government and attracting foreign direct investment, all of which promote expansion of an economy.

There is a need for more and better historical demographic data to be collected on Mexico. Additional measures of well-being must be added. Many of the necessary measures have not been tracked except on a decennial basis or sporadically at best. Some data is entirely missing. Missing data will need to be reconstructed and continued measurement data must be collected. Measurements must focus on population well-being including households and subjective measures as well as market and production data.

Acknowledgments

Special thanks to the faculty of the International Development Program at the University of Southern Mississippi, and especially to Dr. Shahdad Naghshpour whose knowledge, expertise and review of this paper have been invaluable in its development. All errors and omissions are my own.

References

Allee, W.C., A.E. Emerson, Orlando Park, Thomas Park, K.P. Schmidt. 1949. Principles of Animal Ecology. W.B. Saunders Co.

Barbier, Edward B., and Patrik T Hultberg. 2007. "Economic Integration, Environmental Harmonization and Firm Relocation." Environment and Development Economics 12, no. 3 (June 1, 2007): 379. http://www.proquest.com.ezproxy.jbu.edu/ (accessed November 1, 2008).

Beaulieu, Eugene. 2000. "The Canada-U.S. Free Trade Agreement and Labour Market Adjustment in Canada." 2000. The Canadian Journal of Economics / Revue canadienne d'Economique 33, No. 2 (May, 2000), pp. 540-563. Published by: Blackwell Publishing on behalf of the Canadian Economics Association. Article Stable URL: http://www.jstor.org/stable/2667445. (accessed February 28, 2011).

Bhagwati, Jagdish. 2004. In Defense of Globalization. Oxford: Oxford University Press, ix.

Bulmer-Thomas, Victor. 2003. *The Economic History of Latin America since Independence.* Cambridge: Cambridge University Press.

Carlsen, Laura. 2008. "Armoring NAFTA: The Battleground for Mexico's Future." NACLA Report on the Americas 41, no. 5 (September 1, 2008): 17-22, 41. http://www.proquest.com.ezproxy.jbu.edu/ (accessed October 26, 2008).

Delgado-Wise, Raúl, and Humberto Márquez Covarrubias. 2007. "The Reshaping of Mexican Labor Exports under NAFTA: Paradoxes and Challenges." The International Migration Review 41, no. 3 (October 1, 2007): 656-679. http://www.proquest.com.ezproxy.jbu.edu/ (accessed October 26, 2008).

Diep, Kelly. 2008. "Wavering Amigos: The Politics of American Integration." Harvard International Review 30, no. 1 (April 1, 2008): 9-10. http://www.proquest.com.ezproxy.jbu.edu/ (accessed October 26, 2008).

Ederington, Josh. 2007. "NAFTA and the Pollution Haven Hypothesis." Policy Studies Journal 35, no. 2 (May 1, 2007): 239-244. http://www.proquest.com.ezproxy.jbu.edu/ (accessed November 1, 2008).

Fayyaz, Zeina . 2008. "No Easy Peace: The Zapatistas' Tense Stalemate. Harvard International Review 29, no. 4 (January 1, 2008): 9-10. http://www.proquest.com.ezproxy.jbu.edu/ (accessed October 26, 2008).

Hufbauer, Gary Clyde, and Jeffrey J Schott. 2008. "NAFTA's Bad Rap." The International Economy 22, no. 3 (July 1, 2008): 19-23. http://www.proquest.com.ezproxy.jbu.edu/ (accessed October 26, 2008).

Krugman, Paul. 1993. "The Uncomfortable Truth about NAFTA: It's Foreign Policy, Stupid". Foreign Affairs 72, no. 5 (November 1, 1993): 13. http://www.proquest.com.ezproxy.jbu.edu/ (accessed November 1, 2008).

Kuznets, Simon, 1955. "Economic Growth and Income Inequality." The American Economic Review Vol. 45, No. 1 (Mar., 1955): 1-28. Article Stable URL: http://www.jstor.org/stable/1811581
(accessed March 21, 2011).

Nápoles, Pablo Ruiz . 2007. "Protectionism, free trade and preferential trade: the Mexican experience 1970-2005." Banca Nazionale del Lavoro Quarterly Review 60, no. 240 (March 1, 2007): 49-56+. http://www.proquest.com.ezproxy.jbu.edu/ (accessed November 1, 2008).

Piérola, Fernando, and Gary Horlick. 2007. "WTO Dispute Settlement and Dispute Settlement in the "North-South"

Agreements of the Americas: Considerations for Choice of Forum." Journal of World Trade 41, no. 5 (October 1, 2007): 885-908. http://www.proquest.com.ezproxy.jbu.edu/ (accessed October 26, 2008).

Sawyer, W. Charles, and Richard L. Sprinkle, 2004. International Economics, Second Edition. Upper Saddle River: Pearson Prentice Hall.

Shaw, Albany R., 1985. "Recent Developments in NAFTA Law." Law and Business Review of the Americas 14, no. 1 (January 1, 2008): 173-175. http://www.proquest.com.ezproxy.jbu.edu/ (accessed October 26, 2008).

Ravenstein, E.G. 1885. The Laws of Migration, Journal of the Statistical Society of London, Vol. 48, No. 2 (Jun., 1885, 167-235, accessed March 5, 2010.

Trefler, Daniel. 2004. "The Long and Short of the Canada-U.S. Free Trade Agreement." The American Economic Review 94, no. 4 (September 2004): 870-895. http://www.jstor.org/stable/3592797 (accessed March 3, 2011).

Stiglitz, Joseph. 2009. "Towards a better measure of well-being." *Financial Times, FT.com (*Published: September 13 2009). http://www.ft.com/cms/s/0/95b492a8-a095-11de-b9ef-00144feabdc0.html#axzz1GXE2IQ2O (accessed March 2, 2011).

Stiglitz, Joseph, Amartya Sen, and Jean-Paul Fitoussi. 2009. "Report by the Commission on the Measurement of Economic Performance and Social Progress." Commission on the Measurement of Economic Performance and Social Progress (2009) http://www.stiglitz-sen-fitoussi.fr/documents/rapport_anglais.pdf (accessed March 8, 2011).

White, Timothy J. 2007. "NAFTA Revisited: Achievements and Challenges." Review of title_of_work_reviewed_in_italics, clarifying_information. Journal of Third World Studies 24, no. 2 (October 1, 2007): 236-238. http://www.proquest.com.ezproxy.jbu.edu/ (accessed October 26, 2008).

United Nations Educational, Scientific, and Cultural Organization (UNESCO) Institute for Statistics, http://www.indexmundi.com/facts/mexico/, (accessed January 9, 2012).

United States Census Bureau. U.S. Census Bureau - International Programs, International Data Base. http://www.census.gov/population/international/data/idb/country.php, (accessed January 7, 2012).

World Bank Database, Data – WDI, GDF, and ADI Online Databases, http://web.worldbank.org/WBSITE/EXTERNAL/DATASTATISTICS/0,,contentMDK:20398986~menuPK:64133163~pagePK:64133150~piPK:64133175~theSitePK:239419,00.html.

An Empirical Analysis of Assortative Mating Patterns in China

Xu Zhang, Farmingdale State College

Abstract

Economists have been investigating how individuals sort themselves into marriage. Indeed, it not only depends on the characteristics of the alternative future spouses, but also relates to the traits of the choosers. An individual try to find the optimal candidate of future spouse by maximizing individual utility of marriage. This paper examines conditional logit model with data China Health Nutrition Survey (2006) and finds positive assortative mating in all the traits and no evidence of negative assortative mating in labor-market traits. However, the positive assortative mating pattern in the characteristic of hours of work in labor markets is weaker.

Introduction

Marriage patterns in contemporary society have undergone profound changes in the last twenty years. Divorce rate has been rising, and the age at first marriage has increased for both men and women. The transformation in marriage patterns, interacting with rising female labor participation rate and growing educational attainments in population, imply the changes in the characteristics, such as age, wage, earnings, education level, of potential partner when marriage formation is considered. As Becker (1981) addresses who married with whom, "assortative mating" is defined as the similarity (dissimilarity) of the characteristics of husband and wives. The study of "assortative mating" in marriage is important due to the following reasons: first, it may shed some light on the explanation of gender wage gap within marriage; second, it will help to better understand the age gap at first marriage between men and women. one of the payoffs of marriage is having children. Since women have shorter fecundity horizon than men, women tend to marry early in order to maximize their potential utility of marriage, which result in positive age gap at first marriage between men and women.

Previous literature has examined assortative mating in marriage in both biological traits such as age, height, weight and labor market characteristics such as wage, earnings, hours of work. Positive assortative mating between partners is identified when a person matches with one with similar traits while negative assortative mating occurs when a person matches with one with dissimilar traits. Researchers have found some evidences to support both positive assortative mating and negative assortative mating across various types of couples and different countries. For example, Jepsen and Jepsen (2002) documents positive assortative mating are observed for all traits and across all types of couples (same-sex and opposite sex couples) by using 1990 census data, but the positive assortative mating is stronger for non-labor-market traits (e.g. age and education) than for labor-market traits (e.g., hourly earnings). Dalmia and Laurence (2001) find positive assortative mating with respect to age and education in India and the U.S. while no evidence is found to support negative assortative mating in term of labor-market traits, such as wages or hours worked.

Existing literature on assortative mating patterns in China mainly focuses on univariate characteristics of partners within marriage. Mu and Xie (2011) find positive assortative mating in age of marriage appear to be stronger before 1990s while negative assortative mating in age tends to be strengthened during the economic reform-era.

The objective of this paper is to investigate multivariate marriage traits among Chinese people and look for the evidence of positive (or negative) assortative mating patterns as predicted by Becker. The rest of the paper is organized as follows. Section II outlines the literature on economics of marriage which relates to this analysis. Section III introduces the estimation methodology. Section IV describes the data. Section V reports the results. Conclusion and discussion about the results are presented in section VI.

Theory

As Becker's theory of marriage (1973,1974, 1981, 1985) implies, there are several sources of gains from marriage. The first source of gains comes from specialization between household production and labor market activities. A couple will enjoy more satisfaction from marriage if one partner who has comparative advantage in market activities engages in production in the market while the other one focuses on household production. Typically, husbands spend more time on market work, accumulate more human capital and therefore, have comparative advantage in market activities over household production. One important reason of forming a marriage is to have children within the wedlock. Thus, having children, or the production of household public goods becomes the second source of gains. Third, economy of scale in household production

achieves when, for example, food preparation for two is less costly than that to be done separately. Plus, gains from marriage may arise from factors such as tax laws.

Each of these gains from marriage has implication on mating selection process. For instance, to pursue specialization and exchange within the household, division of labor will arise which imply couples tend to be dissimilar in characteristics related with productivity. If having children is important, due to the biological fact that women have shorter fecundity than men, younger women are preferred, a match with positive age gap between husband and wife is therefore implied. As Becker (1974) states an individual maximizes his (her) utility of marriage by selecting a future spouse based on the characteristics of the candidates. "positive (negative) assortative mating" on a characteristic occurs when individuals tend to match with partners who are similar (dissimilar) with respect to that characteristic. Lam(1988) demonstrates positive assortative mating arises when gains from marriage draw from production household public goods (e.g. having children) and negative assortative mating occurs when gains from marriage are due to specialization and exchange within the household. The net effect is ambiguous if both sources contribute to the gains from marriage. Becker (1991) views characteristics such as age, education, race, religion as complements and predicts positive assortative mating for these characteristics; he views characteristics such as wages and hours of work in the labor market as substitutes and predicts negative assortative mating.

A great number of evidence on matching with respect to the husband-wife characteristics has been done, though most of them focus on one characteristic of individuals who form a marriage. In term of age, Hayes (1995) and Over and Phillips (1997) document men prefer younger women and women prefer slightly older men within the marriage outset; while Qian (1998) finds there is a tendency toward homogamous mating with respect to age. Another characteristic that has been widely studied is education. Mare (1991) demonstrates that educational homogamy has increased from the 1930s through the 1980s.

Equilibrium search theory has been applied in the examination of who matches with whom within marriage (e.g. Burdett and Coles, 1997). In this context, marriage is considered as a process where men and women simultaneously search for partners and make a decision comparing gains from marriage with their current utility. And the decision is made based upon characteristics of alternative candidates of partners. Most of these studies are descriptive in nature and do not specify the types of partners in term of characteristics. It is worth studying the interaction of various marriage characteristics when individuals consider forming an union.

While many researchers have been seeking evidence to verify Becker's theory of marriage about assortative mating, little has been done on considering multivariate characteristics of mate matching patterns in China. Plus, marriage in China is not only considered as an union between the two individuals, but also an union of families between bride and groom. When individuals search for potential partners, family background of future spouse may be heavily weighted. Therefore, this paper will also address how much individual's family background connects with the likelihood of marriage.

Methodology

This paper examines whether individuals in contemporary China seek partners with similar (or dissimilar) non-labor market and labor-market characteristics. In order to be consistent with existing literature, Pearson correlation matrix is considered as the first step to study the correlation between different characteristics of individuals within the partnership, although this method only focuses on one-to-one relationship without controlling for the effects of other factors on the match process.

As marriage is a process individuals search over a set of characteristics of potential partners simultaneously to maximize their own utility of marriage, logit model is adopted, to study multivariate characteristics at the same time. To study if actual couples are more similar (or dissimilar) in characteristics within marriage than randomly matched couples, I follow the same methodology in Jepsen and Jepsen (2002), where conditional logit model is considered to study the positive or negative assortative mating in both same-sex partners and opposite- sex partners. Seeking for the impact of multivariate characteristics on the probability of two individuals choosing each other as partners, conditional logit model is considered more appropriate than multinomial logit model in that conditional logit model treats the choices being made based upon the characteristics of the alternative candidates while it is choosers' characteristics determine the likelihood of marriage in multinomial logit model.

Conditional logit model, as described in equation (1), presents the probability that individual i chooses partners j over all the other alternatives to maximize individual i's utility of marriage, which depends on the vector of characteristics of j and how they associate with characteristics of individual i, defined as z_{ij}.

$$P_{ij} = \frac{e^{(Z_{ij}\beta)}}{\sum_{k=1}^{J} e^{[Z_{ik}\beta]}} \quad (1)$$

To apply conditional logit model in mating process, both independent variables and dependent variables need to be defined. Since the mating process depends not only on the characteristics(e.g. age, education, wage rate) of the choosers, but also on the association of characteristics of the two individuals who consider to form a union, the independent variables are defined as the absolute value of the difference in the characteristics of the chooser and the potential mate. The next step of applying conditional logit model is to specify the rejected candidates who are considered as alternatives to the successful candidate who become real couple with the chooser. Since the data only record the real matches, it did not account for who were rejected before the marriage. Therefore, a rejected candidate of spouse is created randomly for each head of household in the data to form a rejected match. Once rejected matches are created, I can compare rejected matches with actual matches and apply conditional logit model. The signs of coefficients which capture the impacts of dissimilarity of characteristics between partners on the likelihood of marriage are predicted as follows: The coefficient tends to be negative for characteristics with positive assortative mating, meaning the more difference on the characteristics with positive assortative mating between the two individuals, the less likely they are actual matches; The coefficient tends to be positive for characteristics with negative assortative mating, meaning the difference on the characteristics with negative assortative mating between the two potential partners will enhance the probability of marriage between them.

Equation (2) shows the specific functional form of the conditional logit model described in Equation (1):

$$m = \beta_1 age + \beta_2 schooling + \beta_3 wage + \beta_4 earning + \beta_5 BMI + \beta_6 height + \varepsilon \quad (2)$$

where 1 is assigned to m if it is an actual couple, otherwise 0 is assigned for rejected matches. All the independent variables appearing in equation (2) are defined as the absolute value of the difference on the characteristic between two individuals. ε is the error term.

Data

The data in this study come from China Health and Nutrition Survey (CHNS) 2006, an international collaborative project between the Carolina Population Center at the University of North Carolina at Chapel Hill and the National Institute of Nutrition and Food Safety at the Chinese Center for Disease Control and Prevention. The survey was conducted as 3-day interview by international team of researchers from various backgrounds. It covers nine provinces from north China to south China which vary substantially in geography, economic development, public resources, and health indicators. The data collect both household data and individual data. Each respondent in the data can be identified through household ID and line number in 2006. Since this study examines the characteristics of couples, only couples currently in the marriage are considered, meaning observations with missing information on marriage are ignored in this study.

Results

Table1 (see appendix) shows the summary statistics of characteristics of husbands who are currently married at overall level and by age groups. There are 3767 observations, among which 180 are under 30 years old. 75% of husbands are under age 60 and the average age of husbands in the sample is around 50.4 years old. Due to data limitation, the age at first marriage only records for women who are older than 52 years old, my study keeps all the husbands who were married at the time of interview. Table 2 demonstrates the summary statistics of characteristics of wives who are currently married by the time of interview. Overall speaking, husbands are older and more educated than wives except for the age group younger than 30 years old, where young women tend to be more educated. In term of labor market characteristics, such as monthly earning (adjusted to Yuan in year 2000), hours of work in labor markets, wage rate, husbands work more hours in labor market, make more money and have higher wage rates than wives.

Using Pearson correlation matrix, I first consider the correlation between the various characteristics within couples. If people choose partners with similar characteristics with respect to non-labor-market traits such as age, education, then we expect positive correlation coefficients for these variables. If couples specialize in home and market production within the families, we expect negative correlation coefficients for earnings, hourly wage rates. The correlations, however, do not control for the effects of other variables on the match. As listed in Table 3, the Pearson correlation coefficients for all the characteristics are positive at 1% significance level, which verify Becker's prediction that non-labor-market characteristics, such as age, education tend to be complements. It does not support the prediction that there is negative correlation in labor-market characteristics such as hours of work.

My further study applies the conditional logit model to capture the impacts of difference in characteristics (in absolute term) between couples on the likelihood of marriage. As shown in Table 4, all the coefficients are negative and significant, indicating that the more different in term of characteristics of age, education, height, earning and hours of work in labor

market between couples, the less likely they will form a marriage. Therefore, the data demonstrates positive assortative mating in all the traits and finds no evidence of negative assortative mating pattern in labor-market traits. However, the positive assortative mating pattern in the characteristic of hours of work in labor markets is weaker. Since the independent variables are in absolute values, it means the directions of difference between couples are treated equally. In order to address how nontraditional matching pattern (that is, wives are older, more educated or making more money than husbands) affect the chances of marriage. I incorporate dummy variables for nontraditional matching patterns in the conditional logit model. As demonstrated in Table 5, nontraditional matching patterns, or hypogamy, with traits of age, education and earning will decrease the likelihood of marriage.

Conclusion and future work

China has experienced profound changes since its open economic reform. The study on assortative mating patterns in China helps to reflect the marriage market of contemporary China. By looking at the data from China Health and Nutrition Survey in 2006, I find positive assortative mating pattern for all the non-labor-market traits (e.g. age, education, height) as well as labor-market traits (e.g. hours of work in labor market, earnings). There is no evidence for negative assortative mating for traits such as hours of work in labor market as described in Becker's theory of marriage. The results are also consistent with the assortative mating pattern in Jepsen and Jepsen (2002). There are a few limitation on current study being done: first, the creation of fake pairs may be improved by finding rejected fake spouses locally instead of nationally; second, collinearity problem may exist among the multiple traits; third, in order to better capture the picture of assortative mating pattern in China, this study can be extended back to 1980s before economic reform was implemented.

Marriage in China is not only considered as an union between the two individuals, but also an union of families between bride and groom. When individuals search for potential partners, family background of future spouse may be heavily weighted. Therefore, future study can address how much individual's family background connects with the likelihood of marriage.

References

Becker, Gary S. 1973. "A Theory of Marriage: Part I." *Journal of Political Economy*, 81:812-846.

Becker, Gary S. 1974. "A Theory of Marriage: Part II." *Journal of Political Economy*, 82:S11-S33.

Becker, Gary S. 1981.1991. *A Treatise on the Family* (Cambridge: Harvard University Press).

Becker, Gary S. 1985. "Human Capital, Effort, and the Sexual Division of Labor." *Journal of Labor Economics*, 81:4, 813-46.

Burdett and Coles. 1997. "Marriage and Class." *The Quarterly Journal of Economics*, 112:1, 141-168.

Dalmia, Sonia and Lawrence, Pareena. 2001. "An Empirical Analysis of Assortative Mating in India and the U.S." *International Advances in Economics Research*, 7:4, November, 443-458.

Hayes, A.F. 1995. "Age Preferences for Same- and Opposite-Sex Partners." *Journal of Homosexuality,* 8:9-21.

Jepsen Lisa K. and Jepsen Christopher. 2002. "An Empirical Analysis of the Matching Patterns of Same-sex and Opposite-sex Couples." *Demography,*39:3, August, 435-453.

Lam, David. 1988. "Marriage Markets and Assortative Mating with Household Public Goods." *Journal of Human Resources*, 23: 462-487.

Mare, Robert. 1991. "Five Decades of Educational Assortative Mating." *American Sociological Review*, 56:15-32

Mu, Zheng and Yu, Xie. 2011."Marital Age Homogamy in China: A Reversal of Trend in the Reform Era?" *Research Report 11-742, Population Studies Center, University of Michigan.*

Over, R. and G. Phillips. 1997. "Differences Between Men and Women in Age Preferences for a Same-Sex Partner." *Behavioral and Brain Sciences*, 20:138-40.

Qian, Zhenchao. 1998. "Changes in Assortative Mating: The Impact of Age and Education, 1970 - 1990". *Demography*, 35:3, 279-92.

Appendix

Table 1. Summary Statistics of married husband.

Variable	All Married Husbands: Mean (S.D.)	Married husband by age group: Mean (S.D.)			
Age	50.4367 (13.2877)	age<=30	30<age<=40	40<age<=50	age>50
Schooling	8.2131 (3.8158)	10.1961 (2.3782)	9.2452 (3.0604)	9.4022 (3.007)	6.9259 (4.1687)
Height(cm)	166.4421 (7.7584)	170.005 (5.900)	167.9572 (8.722)	167.534 (7.7217)	164.8901 (7.1989)
Weight (kg)	65.7276 (12.7701)	67.4817 (12.3913)	67.7499 (12.2287)	67.5937 (12.9949)	63.7458 (12.5973)
Ethnicity (Han)	0.8852 (0.3188)	0.8833 (0.3219)	0.8903 (0.3128)	0.9088 (0.288)	0.8706 (0.3358)
Monthly Earning (in year 2000 Yuan)	607.4993 (859.2289)	469.2781 (233.504)	580.7849 (662.1752)	657.423 (1010.893)	602.1057 (906.6889)
Hours of work last year	416.7425 (240.131)	472.0373 (229.5709)	436.7926 (238.5804)	452.1889 (232.4811)	356.7618 (239.1431)
Wage rate	2.8442 (5.2599)	1.9972 (1.3482)	2.7196 (3.6505)	3.0888 (6.0573)	2.8594 (6.0577)

Table 2: Summary statistics of married wives

Variable	All married wives: Mean (S.D.)	Married wives by age group: Mean (S.D.)			
Age	48.5854 (12.8447)	age<=30	30<age<=40	40<age<=50	age>50
Schooling	6.5975 (4.3174)	9.6354 (2.7979)	8.413 (3.1764)	7.6862 (3.8867)	4.5361 (4.3173)
Height(cm)	155.8577 (7.6099)	158.5502 (5.3149)	157.5709 (6.7634)	156/7629 (7.99)	154.077 (7.658)
Weight (kg)	57.8015 (11.7931)	55.4822 (12.9154)	57.6657 (11.3816)	59.5558 (12.2794)	57.1781 (11.3670)
Ethnicity (Han)	0.8826 (0.3219)	0.8839 (0.3089)	0.8898 (0.3133)	0.9009 (0.2989)	0.867 (0.340)
Monthly Earning (in year 2000 Yuan)	435.6052 (477.0816)	392.6464 (221.9297)	404.6716 (249.8287)	479.2705 (717.9826)	448.6258 (294.7559)
Hours of work last year	380.8592 (248.9084)	441.5829 (265.4836)	413.5845 (245.6136)	401.3914 (240.1255)	300.3966 (238.087)

Table 3: Pearson correlation matrix

Variables	Pearson Correlation Coefficients		
	All	Urban Site	Rural Site
Age	0.9688*	0.9672*	0.9697*
Schooling	0.6417*	0.6575*	0.6099*
Earnings	0.6052*	0.2927*	0.7800*
Hourly Wage	0.2401*	0.2352*	0.3241*
Ethnicity (Han)	0.8531*	0.7921*	0.8670*
Hours of work	0.6134*	0.4306*	0.6250*
Height	0.3081*	0.3789*	0.2719*
weight	0.2516*	0.1847*	0.2835*

*significant at the 1% level, ** significant at the 5% level, *** significant at the 10% level

Table 4: Regression Result of Conditional Logit Model

Variables	Coefficient	Marginal Effects
Diff_Age	-0.7417 ** (.18823)	-2.853
Diff_Schooling	-0.6718** (.28093)	-2.17
Diff_hours of work	-0.01* (.0056)	-1.8
Diff_Earning	-0.0042** (.002)	-2.11
Diff_Height	-0.2277** (.1098)	-2.07

*significant at the 1% level, ** significant at the 5% level, *** significant at the 10% level
Standard errors are shown in parentheses.

Table 5: Regression result from conditional logit model with nontraditional role variables

Variables	Coefficient	Marginal Effects
Diff_Age	-0.34528 ** (.07594)	-4.547
Diff_Schooling	-0.16832 * (.10154)	-1.658
Age_ntr	-1.18563 ** (.59182)	-2.003
Schooling_ntr	-0.61572** (.81879)	-0.752
Earning_ntr	1.17151** (.90021)	1.301

*significant at the 1% level, ** significant at the 5% level, *** significant at the 10% level

When Shareholder Wealth and National Security Interests Collide: An Examination of Export Control Act Violations

Joan Wiggenhorn, Florida Institute of Technology
Kimberly C. Gleason, University of Pittsburgh

Abstract

This paper examines the impact of violations of export policy from a MNE's perspective. We describe the kinds of regulation in place in the U.S. over the past decade and we investigate the response of shareholders to news regarding trade violations and find that they suffer a statistically significant -1.15% abnormal returns over the three day announcement window. We find that the reaction post 9-11 is worse as is the reaction when the export violations occur with countries perceived to be corrupt. That penalty continues in the long run. We also find that both systematic and total risk increase.

Introduction

Following the attacks on the World Trade Center on September 11, 2001, a tremendous amount of reassessment of American foreign policy has taken place by the U.S. citizenry, regulatory institutions, and political participants. National attention has also been drawn to the role of trade as an instrument of political policy. Well known examples of current uses of trade as foreign policy weapons include the U.S. embargos against Cuba, Libya, North Korea, Iran, and the Iraqi Baa'thist regime under Saddam Hussein. While some trade restrictions are directed at individual countries as those just mentioned, others restrict the sale of "high tech" exports to a large number of countries, even allies; and finally, some restrictions apply not just to countries, but to particular individuals.

However, there have been numerous instances in which U.S. firms have been hurt by trade restrictions, particularly when used for foreign policy rather than national security purposes. Export control costs occur due to export losses to companies from blocked contracts, and from foregone future business opportunities, and FDI reductions (Nollen, 1987). One study in 1987 estimated that U.S. exporters lost $9.3 billion in sales annually because of the delays and uncertainty of the export licensing system (Erickson, 1997). Nollen (1987) estimates that the long run and indirect loss to American companies that were involved in the Soviet gas pipeline controls were between $3.6 and $4.1 billion over a six year period.

The fines imposed when companies are found to have violated trade restrictions can be significant. In 2003, Raytheon agreed to pay $25 million in civil fines to settle federal charges that, from 1990 to 1997, it tried to evade export laws in the attempted sale of sensitive radio equipment to Pakistan. In 2007, I.T.T. Corporation pled guilty to multiple criminal charges of the Arms Export Control Act as a result of selling high technology night vision equipment to Indian companies. The fine was a $100 million settlement, and the NV division was prohibited for three years from participating in exporting defense articles or furnishing defense services (Waite and Schwartz, 2007).

In light of both the need for some export controls and the desire of U.S. firms to have the fewest restrictions possible, this paper is the first to examine the impact of export restrictions from an international business perspective. We describe the kinds of regulation in place in the US over the past decade, and identify the characteristics of firms who have been caught and fined for violations of and export restrictions. Furthermore, we investigate the response of shareholders to news regarding trade violations. We find that the impact on firms found violating the EAR can significantly exceed that of the fines alone. On average, the abnormal return over the three day window of the announcement is -1.15%. Additionally, the returns are worse post 9-11 and for violations involving countries that are considered corrupt.

The rest of the paper is organized with the Literature Review first, which includes the literature for both corporate illegalities and export control act regulations. This is followed by the Hypotheses section, and then a description of the firms found in violation of the ECA. We next describe both the event study and long run performance methodology. Following this are the Results and finally the Conclusion and suggestions for further research.

Literature Review

Corporate criminal liability is a frequently discussed topic. While some offenses, such as fraud, theft, tax violations, bribery and antitrust violations have quantifiable fine provisions, there are no such provisions for most environmental, food and drug, and export control violations (Viano and Arnold, 2006). Some such as Khanna (1996) even argue that civil

liability may be just as efficient as criminal liability. However, as Baucus (1989) notes "these questions are of interest and great importance not only to researchers but to lawmakers, managers, and the public in general". Thus, the question remains concerning what is the proper punishment for these export violations.

Some studies have used an event study methodology to determine if the market punishes those firms guilty of criminal activities. In one of the earliest studies, Strachan, Smith and Beedles (1983) report that a firm's stock price is affected negatively on the day before and the day of an announcement of bribery, criminal fraud or certain antitrust activities. Davidson and Worrell (1988), using event study methodology on a sample of announcements of corporate illegalities, find that the market reacts negatively upon announcements of crimes such as bribery, tax evasion, theft of trade secrets, and financial reporting violations. Davidson, Worrell and Lee (1994) examine the stock price reaction to announcements of corporate criminal activity, which they define as "social irresponsibility". They confirm the Davidson and Worrell (1988) results, and furthermore investigate repeat offenders. They find that penalties assessed failed to provide a substantial enough disincentive to deter executives from pursuing illegalities in the future. Karpoff and Lott (1993) examine 132 fraud cases from 1978 through 1987 and find an average decrease of 1.34% for corporate fraud and 5.05% for fraud against government agencies. Karpoff, Lee, and Martin (2011) find large, significant negative abnormal returns to firms prosecuted for foreign bribery. However, none of these studies investigated corporate crime in the form of export control violations.

In a meta-analysis, Frooman (1997) evaluates 27 prior event studies that measured the stock market's reaction to incidences of socially irresponsible and illicit behavior. He finds that shareholder wealth decreases are statistically significant. He also finds that for the sample for illegal behavior is more negative than that for what he calls "irresponsible behavior". Two studies illustrate that the negative reaction may be dependent on how the firm reacts to the violation announcement. In examining product safety liability lawsuits, Viscusi and Hersch (1990) find that the negative stock market reaction is greater the longer the newspaper coverage. However, Haslem (2005) finds that companies are punished more when they settle than if they continue the litigation process until a judgment is received. Thus it is unclear if firms should admit guilt quickly and settle or allow the process to unfold.

The US Department of Commerce's Bureau of Industry and Security (BIS) handles enforcement of import-export restrictions for reasons of national security, foreign policy, and nonproliferation – falling under the category of Export Administration Regulations (EARs). Investigation and prosecution of potential violators is conducted by the US Department of Justice, with the support of the BIS Export Enforcement Special Agents unit. The BIS also provides a database of regulations, as well as a list of names of "Denied Persons" with whom American entities cannot engage in import-export activities. Managers of international businesses who plan to engage in import-export activities are responsible for putting controls in place to ensure that the hosts of restrictions are not violated. Morris (2006) stresses the importance of executives in culturing export control compliance. Penalties are also dependent on eight aggravating factors and nine mitigating factors. Of these nine, the two given greatest weight are voluntary self disclosure of the violation and the company's having an effective export compliance culture.

The primary kind of restriction is prohibitions on certain goods, such as high-tech items chemicals that could be used in the development of weapons. For example, the US Chemical Weapons Convention (CWC) prohibits firms worldwide from exporting a wide range of substances. The CWC EAR regulations are quite complex, and compliance – recordkeeping, reporting, and disclosure on behalf of exporting companies is expensive and cumbersome. Firms violating EARs are subject to both criminal and administrative prosecution, and penalties ranging from fines varying with the severity of the violation to revocation of export privileges. Another example are the multilateral export control agreements in which the US participates, such as the Wassenaar Arrangement on Export Controls for Conventional Arms and Dual Use Goods and Technologies, the Missle Technology Control Regime, the Nuclear Suppliers Group, and the Australia Group, which focuses on chemical weapons precursor exports.

Exports with "denied parties" are a second category of criminal offenses related to trade. The BIS maintains its list of denied persons, as well as an EAR Entity List, of organizations suspected of facilitating the proliferation of weapons of mass destruction; the US Treasury has a Specially Designated Nationals and Blocked Persons List, consisting of persons suspected of narcotics trade or terrorist financing. In "Fines for the Unsuspecting", Rogers (2006) warns that it is important to know with whom you are doing business. For instance, all Cubans aren't listed individually on the SDN list, but a transaction with any Cuban national is prohibited. Violations of U.S. or U.N. embargoes are prosecutable under law; exports and re-exports to and from Iran, Iraq, Afghanistan, Angola, Libya, North Korea, Rwanda, Sudan, Syria, and Yugoslavia are severely restricted, if not prohibited. The Treasury maintains a separate list of countries, including Burma, the Balkans, Liberia, and Zimbabwe, and industries and activities, such as diamond mining, narcotics trafficking, proliferation of nuclear weapons, under sanction or forbidden by US regulation. The denied persons, entity, embargoed, and sanctioned lists are all updated as information regarding threat potential or human rights situations changes. Jeffrey Bialos (1989) highlights some of the difficulties in the detection and disclosure of violations of U.S. export controls. Since the process depends on voluntary disclosure by a firm, when a company learns that it may have violated U.S. export controls it must first decide what type of internal investigation

to conduct and what processes to adopt. Initial remedial measures can mitigate potential sanctions and demonstrate a commitment to compliance with U.S. export control laws.

We examine a sample of 75 export violations. If we assume that criminal activity conducted by the firm or employees on behalf of the firm is profitable, then criminal activity should increase owners' wealth as long as the crime goes undetected. However, when the criminal activity has been discovered, the firm may face several costs including fines, legal fees, lost employee time (spent fighting the charges), and negative publicity. While the criminal activity may have been profitable, "getting caught" will certainly not be and should result in a negative stock price reaction.

Hypotheses

Given that most studies (Davidson and Worrell, 1988; Karpoff and Lott, 1993; Davidson,Worrell, and Lee, 1994) find a negative reaction to announcements of corporate illegality, we expect:

H1: The announcement of export violations by corporations will be accompanied by negative abnormal returns.

There also appears to be factors which either strengthen or mitigate the negative announcement returns. Karpoff and Lott (1993) find that the announcement of alleged or actual corporate fraud results in a statistically and economically significant loss in the market value of the common stock. Karpoff, Lee, and Martin (2011) find that when charges of financial misconduct in addition to bribery are reported, the share price response is especially pronounced. Block (1991) argues that penalties assessed against a corporation for detected violations are important because they affect a firm's future calculations for possible violations. Baucus and Baucus (1997) find that stock prices react strongly to multiple convictions. They also find that the more serious the offense, the more negative the long term performance. Some papers have found a relationship between MNE's returns and joint ventures with foreign governments based on the level of corruption (Gleason et al, 2005). Thus, since a stigma effect has been found in previous studies, it would seem to follow that:

H2: Announcement effects will be more negative for violations with countries that are viewed as corrupt.

Karpoff and Lott (1993) argue that the "reputational cost of corporate fraud is large and constitutes most of the cost incurred by firms accused or convicted of fraud (p. 758). These reputational effects may be long-lasting. Baucus and Baucus (1997) also examine the longer term performance effects of corporate illegality and show that firms experience lower accounting returns over five years and slower sales growth in the third through the fifth year. Although they do not find significant poorer long run stock performance, part of the problem may be in measuring long run abnormal returns. Hence, acknowledging the difficulties in measurement, we still hypothesize that:

H3: Firms with export violations will suffer long term performance effects.

The last consideration is whether or not firms that have violated the ECA will then have greater systematic risk. Given the unfavorable publicity and investors' possible reaction to the news, it is possible that these firms will then have greater systematic risk.

H4: Firms with export violations will experience an increase in both total risk and systematic risk.

Data and Methodology

Sample Selection

The study's data set was initially taken from a data base search of Lexis Nexis Newsires and the US Department of Commerce for the years 1980-2010. For a firm to be included in the study, its stock had to be listed on the New York, or NASDAQ Stock Exchanges. This criterion serves two purposes. First, it ensures that the stock is traded frequently enough for market model parameters to be estimated. Second, we require the firm to have returns data available from the Center for Research in Security Prices (CRSP). Accounting data is obtained from Standard and Poor's Research Insight.

The original data set of violations of import export restrictions was obtained from the U.S. Department of Commerce Bureau of Industry and Security. To identify the addition details of the violation, newswire releases for each incident of illegality was obtained illegality were obtained from Lexis Nexis and examined to discern the key variables surrounding both the crime event and the outcome and penalty.

There were ECA violations over the entire 16 year period with 44% occurring prior to 9-11-2001. However, nearly half of all the violations occurred during the four year period from 2004 through 2007. Concerning the sic codes, almost half of the violations were for sic codes in the 3000s. This is not surprising since these SIC codes include fabricated metal products, industrial machinery and equipment, electrical and electronic equipment and instruments and related products that include search, detection, navigation, guidance, aeronautical, and nautical systems and instruments. Regarding the characteristics of the offending companies, while the asset size varies, the mean and median market value of the firms involved are about $28

million dollars. They tend to be profitable as both the ROA and ROE attest, but there is considerable variation in the ROE percentages.

Event Study Analysis

The major purpose of this study is to determine the stock market's reaction to the announcements of Export Control Restrictions. Event study methodology is used to identify the wealth effects to bidders subsequent to the announcements of acquisitions of family owned targets. Daily excess (abnormal) returns are calculated by subtracting the expected returns from the actual returns for each day of the window in question. The ordinary least squares market model is used to specify the expected returns generating process. We use the market model parameter over the estimation period from t = -110 to t = -11 relative to the announcement day t=0. The standardized cross-sectional method Boehmer, Musumeci, and Poulsen (1991) with Scholes-Williams (1977) betas is used to test for significance.

For each firm, the abnormal return is calculated as the difference between the firm's actual return and the expected return generated by the market model. Thus, the abnormal returns are calculated as:

$$AR_{ij} = R_{ij} - ER_{ij} \quad (1)$$

Where AR_{ij} is the abnormal return on day i for each firm j; R_{ij} is the actual return on day i for each firm j, and ER_{ij} is the expected return generated from the market model on day i for each firm j.The average excess return for any day is calculated by summing over the ARs for the N firms in the sample and dividing by N. The cumulative average excess returns (CARs) over a multi-day event period are calculated by summing the average excess returns over the T day event window.

Analysis of Long Run Returns

We further investigate the post-conviction performance of firms violating the Export Control Act. The long run stock performance is examined using the Fama-French 3 factor model. The BHAR methodology suggested Lyon, Barber, and Tsai (1999) has been found to suffer from the cross correlation problem arising because matching on firm-specific characteristics cannot completely remove the correlation between event firms' returns. More recently, long run performance has been measured using the Fama-French 3 factor model (Yook, 2010; Hyland, 2008; Ferreira, Sinha, Varble, 2008). Yook (2010) suggests that the BHAR methodology presents a modeling problem since "no benchmark provides a perfect estimate of the expected return and errors from the difference between the return of an event firm and its benchmark are compounded in computing long-term buy-and-hold returns" (Yook, 2010, 325-326). The returns are calculated using the Fama-French three factor model using the formula:

$$R_{p,t} - R_{f,t} = \alpha + \beta(R_{m,t} - R_{f,t}) + sSMB_t + hHML_t + \varepsilon_{p,t} \quad (2)$$

Where $R_{p,t}$ is the event portfolio's return in month t ; $R_{f,t}$ is the 1 month Treasure bill rate, observed at the beginning of the month; $R_{m,t}$ is the monthly market return in month t; SMT_t is the average return on small market capitalization portfolio minus the average return on a large market-capitalization portfolio; HML_t is the average monthly return on a high book-to-market equity portfolio minus the average monthly return on a low book to market equity portfolio. (Fama and French, 1993).

Risk Assessment

Any investigation into the effects of export violations should include whether the subsequent reactions lead to higher risk to shareholders. We provide a perspective on risk, using an empirical application of portfolio theory. Systematic risk, or *beta*, measures the covariance of an asset's returns with that of the market. It can also be written as:

$$\beta_i = \rho_{i,m}\, \sigma_i / \sigma_m \quad (3)$$

Where $\rho_{i,m}$ is the correlation coefficient between the return for firm i and the market; σ_i is the standard deviation of firm i's returns, and σ_m is the standard deviation for the market's. Systematic risk as well as total risk as measured by the standard deviation are computed both before and after the determination of the export violation and the change measured.

Results

Event Study Results

Table 1: Short Run Returns

	Number	Period (0, 0)	T stat	Period (-1, +1)	T stat	Period (-1, 0)	T stat
Full Sample							
Full	75	-0.41	(-1.14)	-1.15	(-2.73)***	-0.65	(-1.11)
Pre 9/11	33	0.13	(0.51)	-0.05	(-1.06)	-0.52	(-1.22)
Post 9/11	42	-0.83	(-2.19)**	-1.28	(-2.02)**	-1.21	(-1.85)*
T-test (Post–Pre)			(-1.88)*		(-2.27)**		(-1.55)
Hi Corrupt	46	-2.23	(-2.57)**	-2.71	(-2.04)**	-2.66	(-2.27)**
Lo Corrupt	29	0.70	(1.42)	0.10	(1.04)	0.34	(1.38)*
T-test (Hi–Lo)			(-2.65)**		(-3.09)***		(-2.16)**

*indicates 10% significance; ** indicates 5% significance *** indicates 1% significance*

The results in Table 1 indicate that violations of the export control act, overall, are viewed negatively by the market, resulting in a 1.15% loss of shareholder value over the (-1,+1) event window, significant at the 1% level. While the results for pre 9/11 are insignificant, post 9/11 are negative and significantly different from the pre 9/11 results. We also run tests on the returns based on the corruption perception of the foreign firm involved. Here, as one might expect, the returns are over -2% for all three windows for the high corruption countries and in general insignificant or even positive over the two day window. However, the difference between the means for the two groups of countries is statistically significant.

Long Run Performance

The long run results are presented in Table 2. They strongly indicate that firms that violate the Export Control Act are not penalized by the market in the long run. Except for the six month window, all the results are insignificant, though positive. These results suggest that there is no penalty in terms of long term performance of violators, supporting the "business as usual" conjecture. The pre 9/11 and post 9/11 results are even more puzzling. In the post 9/11 period results are significantly higher than during the pre-9/11 period. Concerning the High Corruption versus Low Corruption results, high corruption is significantly negative, as expected, but the results for Low Corruption are remarkably high. Apparently a year to 2 years after the violation, if the country involved was not seen to be corrupt with no apparent security risk, than the company involves experiences positive returns.

Table 2: Long Run Returns

	Period (0,6)	T stat	Period (0,12)	T stat	Period (0,18)	T stat	Period (0,24)	T stat
ECA (N=72)	6.72	(2.17)**	2.13	(0.49)	3.65	(0.69)	5.64	(0.91)
Pre-9/11 (N= 31)	5.58	(1.25)	-1.03	(-0.16)	2.24	(0.29)	2.23	(0.25)
Post 9/11 (N=41)	7.60	(1.82)*	4.58	(0.69)	4.75	(0.40)	8.29	(0.89)
T-test (Post–Pre)		(1.25)		(2.26)**		(1.50)		(1.74*)
Hi Corrupt (N=46)	-2.73	(-0.37)	-6.74	(-1.17)	-9.83	(-1.95)*	-15.58	(-2.14)**
Lo Corrupt (N=29)	9.01	(1.63)	10.43	(2.00)**	12.96	(2.59)***	27.69	(2.74)***
T-test (Hi–Lo)		(-1.26)		(-1.40)		(-2.27**)		(-3.05)***

** indicates 10% significance; ** indicates 5% significance *** indicates 1% significance*

Risk Shifts

The last tests which are conducted involve changes in risk. The results presented in Table 3are as expected. Those firms that violate the ECA experience both higher total risk as measured by the standard deviation and higher systematic risk as measured by Beta. The differences between the before and after risk measures were both significant.

Table 3: Risk Results

	Change in Total Risk	**Change in Systematic Risk**
ECA firms	0.009 (4.22)***	0.114 (3.57)***

** indicates 10% significance; ** indicates 5% significance *** indicates 1% significance*

Discussion and Conclusions

The results shown here support the findings of Davidson and Worrell (1988), Karpoff and Lott (1993), Davidson, Worrell and Lee (1994), Frooman (1997), and Karpoff et al (2011) in that the market does react negatively to the release of news of firm illegality. This study also lends support to the notion that the stock market penalty is immediate and that the market reacts efficiently to these announcements. What we have confirmed so far is that both investors and firm officials appear to share the punishment in the short run for such actions. Although some risk-seeking investors may consider it desirable for managers to occasionally get caught trying to cheat, the present results indicate that a marketplace penalty will be imposed. Key executives will have to balance these costs against expected benefits from such irresponsible corporate actions.

We must, however, point out a potential caveat. This study pertains only to whether sensitive investors are influenced by the bad news. We cannot determine if this is motivated by the falling stock price or by ethical consideration. Future research may want to address this issue. There are several issues which merit additional study. Godfrey, Merril and Hansen (2009) find that certain types of CSR create goodwill which lessens the impact of negative legal/regulatory actions. It would be useful to find if firms that are guilty of export violations can also mitigate the negative announcement returns. This could be examined by comparing firms in the top percentiles of the *Fortune* list with those in the bottom percentiles. This issue seems particularly relevant given our long term results that firms that sold to countries with low corruption did well a year after the violation. Additionally, one can study if violations involving national security pay a heavier price than those ad hoc export controls supporting temporary foreign policy. This would require looking more closely at the types of exports that were sold by each company.

References

Baucus, M. S.: 1988, `Who Commits Corporate Wrongdoing: Predicting Illegal Corporate Behavior Using Event History Analysis', *Academy of Management Best Papers Proceedings, pp.* 160-164.

Baucus, M. and D. Baucus:1997,"Paying the Piper:An Empirical Examination of LongerTerm Financial consequences of illegal Corporate Behavior", *Academy of Management Journal,*40,1,129-151.

Bialos, J.: 1989, "The Detection and Disclosure of Violations of U.S. Export Controls: Management Choices and their National Security Implications", 29, 567-620.

Block, M. 1991, "Optimal Penalties, Criminal law and the Control of Corporate Behavior", *Boston University Law Review,* 71, 395-419

Davidson III, W. and D. Worrell: 1988, "The Impact of Announcements of Corporate Illegalities on Shareholder Returns", *The Academy of Management Journal*, 31,1, 195-200.

Davidson III, W., D. Worrell, and C. Lee: 1994, "Stock Market Reactions to Announced Corporate Illegalities", *Journal of Business Ethics*, 13, 979-987.

"Don't Let This Happen to You!" 2008, U.S. Department of Commerce, Bureau of Industry and Security, Export Enforcement. Accessed January 9, 2010 http://www.bis.doc.gov/complianceandenforcement/dontletthishappentoyou-2008.pdf

Erickson, G.:1997, "Export Controls:Marketing Implications of Public Policy Choices", *Journal of Public Policy & Marketing*, 16,1, 83-92.

Frooman J.: 1997, "Socially Irresponsible and Illegal Behavior and Shareholder Wealth", *Business and Society,* 36,3, 221-249.

Gleason, K., Malgwi, C., Mathur, I. and V. Owhoso. (2005). Impact of perceived national

corruption on the returns to U.S. multinationals in transactions with foreign governments, *Review of Accounting and Finance*, Vol. 4 No. 2, pp. 26-51.

Godfrey, P., C. Merrill, and J. Hansen: 2008, "The Relationship between Corporate Social Responsibility and Shareholder Value: An Empirical test of the Risk Management Hypothesis", *Strategic Management Journal,* 30, 425-445.

Haslem, B.: 2005, "Managerial Opportunism during Corporate Litigation", *Journal of Finance,* 60,4, 2013-2041.

"Introduction to the Commerce Department's Export Controls", March 2007.Found at www.bis.doc.gov

Karpoff, J. and J. Lott. Jr.: 1993, "The Reputational Penalty Firms Bear from Committing Criminal Fraud", *Journal of Law & Eocnomics,*36, 757-802.

Karpoff, J., D. Lee, and G. Martin, 2011, "Bribery penalties sting less than penalties for other financial misconduct", Foster School of Business, University of Washington. http://www.foster.washington.edu/centers/facultyresearch/Pages/bribery.aspx

Kerber, R.: 2003, "Raytheon Co. Assessed One of Largest-Ever Penalties for Export Violations, *Knight Ridder Tribune Business News,* February 28, 2003, p. 1.

Khanna, V. :1996, "Corporate Criminal Liability: What Purpose Does It Serve?", *Harvard Law Review*, 109,7,1477-1534.

Khanna, V.:2000,"Corporate Liability Standards:When Should Corporations be Held Criminally Liable?", *The American Criminal Law Review,*37,4, 1239-1283.

Lindsay, J.:1986, "Trade Sanctions as Policy Instruments: A Re-Examination", *International Studies Quarterly,*30,2, 153-173.

Milhollin, G.:2002, "Export Controls: Why We Need Them More than Ever", keynote address before the Practicing Law Institute, December 9, 2002. Accessed on January8, 2010 http://www.wisconsinproject.org/pubs/speeches/2002/12-9-02ecs.htm

Morris, M.: 2006, "The Executive Role in Culturing Export Control Compliance", *Michigan Law Review,*104,7,1785-1805.

Nollen, S.:1987,"Business Costs and business Policy for Export controls", *Journal of International Business Studies*, 18,1, 1-18.

Rogers, R.:2006, "Fines for the Unsuspecting", *Strategic Finance,* 87,12,51-55.

"Senate Eases Export Limits on Some High-Tech Items", 2001, *Wall Street Journal,* September 7, 2001, A8.

Slocum, J.:1997, "Export Controls: Agencies Should Assess Vulnerabilities and Improve Guidance for Protecting Export-Controlled Information at Universities", *The Journal of Research Administration,*38, 119-121.

Viano, M. and J. Arnold: 2006, "Corporate Criminal Liability", *The American Criminal Law Review*, 43,2,311-339.

Viscusi, w. and J. Hersch: 1990, :The Market Response to Product Safety Litigation", *Journal of Regulatory Economics*, 2, 215-230.

Waite, J. and J. Schwartz:2007, "Enforcement Activity Mounts as Department of Justice Focuses on Exporters", *Intellectual Property & Technology,* 19,8,14-18.

Mean Reversion or Random Walk in the Stock Prices of Transition Countries? Sequential Panel Selection Method

Tsangyao Chang and Chia-Hao Lee, Feng Chia University
Ken Hung, Texas A&M International University

Abstract

This study applies the Sequential Panel Selection Method (SPSM) procedure proposed by Chortareas and Kapetanios (2009) to investigate the time-series properties of stock prices for the nine transition countries during the 2000.10 to 2010.11 period. SPSM classifies the whole panel into a group of stationary series and a group of non-stationary series. In doing so, we can clearly identify how many and which series in the panel are stationary processes. Empirical results form the SPSM using the Panel KSS unit root test (Ucar and Omay, 2009), with a Fourier function indicate that the null hypothesis of I(1) unit root in stock prices cannot be rejected for all the transition countries under study.

Introduction

Researchers in finance have long been interested in the time-series properties of equity prices, with particular concern regarding whether stock prices can be described as random walk (unit root) or mean reverting (trend stationary) processes. Whether or not stock prices are characterized by a unit root has important implications for the efficient market hypothesis, which asserts that returns of a stock market are unpredictable from previous price changes. If stock prices are an I(0) stationary process, then any shock effect is temporary. Thus, shifting the stock price from one level to another will eventually return it to its equilibrium level. From an investment point of view, this ensures that one can forecast future movements in stock prices based on past behavior and trading strategies can be developed so as to earn abnormal returns. However, if it is found that stock prices are non-stationary (or I(1) process) then shocks will have a permanent effect, implying that stock prices will attain a new equilibrium and future returns cannot be predicted based on historical movements in stock prices. This proposition was termed the weak-form efficient market hypothesis (hereafter, weak-form EMH). A non-stationary stock price also implies that volatility in stock markets will increase in the long run without bound (Narayan, 2008). Nelson and Plosser (1982) pointed out that whether stock prices are modeled as a trend stationary or as a difference stationary process has important implications vis-à-vis modeling, testing, and forecasting. Studies on this issue are of considerable concern to researchers conducting empirical studies and investors alike.

Recently, there is a growing consensus that stock price exhibits nonlinearities and, consequently, conventional unit root tests, such as the Augmented Dickey Fuller (ADF) test, have low power in detecting mean reversion of stock prices. It is important to note, nevertheless, that under no circumstance does the finding of non-linear adjustment necessarily signify the existence of non-linear mean reversion or stationarity. Thus, it is essential that stationary tests based on a non-linear framework be applied.

More recently, it has been reported that conventional unit root tests not only fail to consider information across regions, thereby leading to less efficient estimations, but also have lower power when compared with near-unit-root but stationary alternatives (Taylor and Sarno, 1998; Maddala and Wu, 1999; Levin *et al.*, 2002; Im *et al.*, 2003; Choi and Chue, 2007; Pesaran, 2007). It is not surprising that these factors have induced considerable doubt on many of the earlier findings, which based on a unit root in stock prices. In order to increase the power in testing for a unit root, many researchers have employed panel data. Levin et al. (2002) and Im et al. (2003), for instance, have developed the asymptotic theory and the finite-sample properties of ADF tests for use with panel data. These two tests have significantly improved power even in relatively small panels. Taylor and Sarno (1998) and Breuer et al. (2001) have shown that the "all-or-nothing" nature of the tests has not been fully addressed by recent methodological refinements to the Levin-Lin-Chu (2002) test. Although Im et al. (2003), Maddala and Wu (1999) and Taylor and Sarno (1998) developed tests that permit the autoregressive parameters to differ across panel members under the stationary alternative, they are not informative in terms of the number of series that are stationary processes when the null hypothesis is rejected. The reason is simple: they are not joint tests of the null hypothesis. In this regards, Breuer et al. (2001) claim that, by analogy to a simple regression, when an F-statistic rejects the null that a vector of coefficients is equal to zero, it is not necessarily true that each coefficient is nonzero. Likewise, when the unit-root null hypothesis is rejected, it may very well not be justified to assume that all series in the panel are stationary. In contrast to those panel-based unit root tests that are joint tests of a unit root for all members of a panel and that are incapable of determining the mix of I(0) and I(1) series in a panel setting, the Sequential Panel Selection Method (hereafter, SPSM),

proposed by Chortareas and Kapetanios (2009), classifies a whole panel into a group of stationary series and a group of non-stationary series. In doing so, we can clearly identify how many and which series in the panel are stationary processes

Perron (1989) argued that if there is a structural break, the power to reject a unit root decreases when the stationary alternative is true and the structural break is ignored. Meanwhile, structural changes present in the data generating process, but have been neglected, sway the analysis toward accepting the null hypothesis of a unit root. The general method to account for breaks is to approximate them using dummy variables. However, this approach has several undesirable consequences. First, one has to know the exact number and location of the breaks. These are not usually known and therefore need to be estimated. This in turn introduces an undesirable pre-selection bias (see Maddala and Kim, 1998). Second, current available tests account only for one to two breaks. Third, the use of dummies suggests sharp and sudden changes in the trend or level. However, for low frequency data it is more likely that structural changes take the form of large swings which cannot be captured well using only dummies. Breaks should therefore be approximated as smooth and gradual processes (see Leybourne et al., 1998). These arguments motivate the use of a recently developed set of unit root and stationary tests that avoid this problem. Both Becker et al. (2004, 2006) and Enders and Lee (2009) develop tests which model any structural break of an unknown form as a smooth process via means of Flexible Fourier transforms. Several authors, including Gallant (1981), Becker et al. (2004) and Enders and Lee (2009), and Pascalau (2010), show that a Fourier approximation can often capture the behavior of an unknown function even if the function itself is not periodic. The authors argue that their testing framework requires only the specification of the proper frequency in the estimating equations. By reducing the number of estimated parameters, they ensure the tests have good size and power irrespective of the time or shape of the break. Additionally, the existence of structure changes in stock prices might imply broken deterministic time trends and the result is a nonlinear pattern (Bierens, 1997). It should, therefore, not be unexpected that these shortcomings have seriously called into question many of the earlier findings based on a unit root in stock prices.

There are several novelties from our study. First, to the best of our knowledge, this study is the first of its kind to utilize the Panel KSS unit root test with a Fourier function through the SPSM procedure to examine evidence for random walk (and/or mean reversion) in stock prices for nine transition countries (i.e., Bulgaria, the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Romania, and Russia). We found the transition countries to be an interesting sample to investigate stock market behavior because they had moved from centrally planned economies towards market driven economies recently. Secondly, it is well-known that independence is not a realistic assumption in that the stock prices of different countries may be contemporaneously correlated. To control for any cross-section dependence found among the data sets, we approximate the bootstrap distribution of the tests and this is not done in the previous study which assume the individuals are cross-section independent. O'Connell (1998) has in fact shown that the true size of both tests can be far greater than the normal size when the underlying data-generating process (DGP) is characterized by cross-section dependence. With these, the current research hopes to fill the existing gap in the literature. Our empirical results indicate that the unit root process is a characteristic of the stock price in these nine transition countries under study.

The remainder of this study is organized as follows. Section 2 presents the data used. Section 3 first describes the methodology employed and then discusses the empirical findings and policy implications. Section 4 presents a wrap up of the conclusions we draw.

Data

The data set consists of weekly stock market indices for nine transition countries: Bulgaria, the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Romania, and Russia. The stock market indices for these nine transition countries are the Sofia Stock Exchange Index for Bulgaria, Prague Stock Exchange Index for the Czech Republic, the OMX Tallin Stock Exchange Index for Estonia, the Hungary-DS Market - Price Index for Hungary, the OMX Riga Stock Exchange Index for Latvia, OMX Vilnius Stock Exchange Index for Lithuania, Poland-DS Market – Price Index for Poland, Romania Bet (L) – Price Index for Romania, and Russia RTS Index – Price Index for Russia. Sample periods cover from October 2000 to November 2010. Table 1 reports the summary statistics of the data studied. We find that Poland had the highest average stock market returns of -0.07% and both the Czech Republic and Hungary have the lowest average stock market returns of -0.12% over this sample period. The measures for skewness and excess kurtosis show that the stock market return series are highly leptokurtic and negatively skewed with respect to the normal distribution, indicating that all stock market returns are non-normal. This result is consistent with that of the current literature.

Sequential Panel Selection Method (SPSM) and Panel KSS Unit Root Test with a Fourier Function

As we stated earlier that there is a growing consensus that stock price exhibits nonlinearities and, consequently, conventional unit root tests, such as the Augmented Dickey Fuller (ADF) test, have low power in detecting mean reversion of stock price. A number of studies have provided empirical evidence on the nonlinear adjustment of exchange rate. However, the finding of nonlinear adjustment does not necessarily imply nonlinear mean reversion (stationarity). As such, stationarity tests based on a nonlinear framework must be applied. Ucar and Omay (2009) proposed a nonlinear panel unit root test by combining the nonlinear framework in Kapetanios et al. (2003, KSS) with the panel unit root testing procedure of Im et al. (2003), which has been proven to be useful in testing the mean reversion of stock price. Perron (1989) argued that if there is a structural break, the power to reject a unit root decreases when the stationary alternative is true and the structural break is ignored. Meanwhile, structural changes present in the data generating process, but have been neglected, sway the analysis toward accepting the null hypothesis of a unit root. Therefore, the Sequential Panel Selection Method (SPSM) proposed by Chortareas and Kapetanios (2009), mixed with the Panel KSS unit root tests with a Fourier function, were used to investigate the time-series properties of stock prices for the nine transition countries.

In line with Kapetanios et al. (2003), the KSS unit root test is based on detecting the presence of non-stationarity against a nonlinear but globally stationary exponential smooth transition autoregressive (hereafter, ESTAR) process. The model is given by

$$\Delta X_t = \gamma X_{t-1}\{1-\exp(-\theta X_{t-1}^2)\} + \nu_t, \tag{1}$$

where X_t is the data series of interest, ν_t is an i.i.d. error with zero mean and constant variance, and $\theta \geq 0$ is the transition parameter of the ESTAR model and governs the speed of transition. Under the null hypothesis X_t follows a linear unit root process, but X_t follows a nonlinear stationary ESTAR process under the alternative. One shortcoming of this framework is that the parameter γ is not identified under the null hypothesis. Kapetanios et al. (2003) have used a first-order Taylor series approximation for $\{1-\exp(-\theta X_{t-1}^2)\}$ under the null hypothesis $\theta = 0$ and have then approximated equation [1] by using the following auxiliary regression:

$$\Delta X_t = \xi + \delta X_{t-1}^3 + \sum_{i=1}^{k} \theta_i \Delta X_{t-i} + \nu_t,\ t = 1,2,....,T \tag{2}$$

In this framework the null hypothesis and alternative hypotheses are expressed as $\delta = 0$ (non-stationarity) against $\delta < 0$ (non-linear ESTAR stationarity). Furthermore, Ucar and Omay (2009) have expanded a nonlinear panel data unit root test based on regression [1]. The regression is:

$$\Delta X_{i,t} = \gamma_i X_{i,t-1}\{1-\exp(-\theta_i X_{i,t-1}^2)\} + \nu_{i,t} \tag{3}$$

Ucar and Omay (2009) have also applied first-order Taylor series approximation to the Panel ESTAR [3] model around θ_i=0 for all *i*, and have obtained the auxiliary regression:

$$\Delta X_{i,t} = \xi_i + \delta_i X_{i,t-1}^3 + \sum_{j=1}^{k} \theta_{i,j} \Delta X_{i,t-j} + \nu_{i,t} \tag{4}$$

Where $\delta_i = \theta_i \gamma_i$ and the hypotheses established for unit root testing based on regression [4] are as follows:

$$\begin{aligned} H_0 &: \delta_i = 0, \text{ for all i, (linear nonstationarity)} \\ H_0 &: \delta_i < 0, \text{ for some i, (nonlinear stationarity)} \end{aligned} \tag{5}$$

Furthermore, the system of the KSS equations with a Fourier function that we estimate is:

$$\Delta X_{i,t} = \xi_i + \delta_i X^3{}_{i,t-1} + \sum_{j=1}^{k1} \theta_{i,j} \Delta X_{i,t-j} + a_{i,1} \sin(\frac{2\pi kt}{T}) + b_{i,1} \cos(\frac{2\pi kt}{T}) + \varepsilon_{i,t} \quad (6)$$

Where $t = 1,2,....,T$. The rational for selecting $[\sin(2\pi kt/T), \cos(2\pi kt/T)]$ is based on the fact that a Fourier expression is capable of approximating absolutely integrable functions to any desired degree of accuracy. Where k represents the frequency selected for the approximation, and $[a_i, b_j]'$ measures the amplitude and displacement of the frequency component. It also follows that at least one frequency component must be present if there is a structural break. Gallant (1981), Becker et al. (2004) and Enders and Lee (2009), and Pascalau (2010), show that a Fourier approximation can often capture the behavior of an unknown function even if the function itself is not periodic. As there is no a priori knowledge concerning the shape of the breaks in the data, a grid-search is first performed to find the best frequency.

The SPSM proposed by Chortareas and Kapetanios (2009) are based on the following steps:

(1) The Panel KSS test with a Fourier function is first conducted to all real exchange rates in the panel. If the unit-root null cannot be rejected, the procedure is stopped. Therefore, all the series in the panel are nonstationary. If the null is rejected, go to Step 2.
(2) Remove the series with the minimum KSS statistic since it is identified as being stationary.
(3) Return to Step 1 for the remaining series or stop the procedure if all the series are removed from the panel.

Final result is a separation of the whole panel into a set of mean-reverting series and a set of non-stationary series.

Empirical Results, Economic and Policy Implications

Unit Root Tests

Table 2 presents the country-by-country for the unit root and stationary tests (i.e., the ADF; P-P, 1988; and the KPSS, 1992). At first sight, the individual unit test statistics seem to show that stock prices are non-stationarity for all nine transition countries. As stated earlier, there is a growing consensus that conventional unit root tests such as the ADF and PP tests - fail to incorporate the structural breaks in the model - have low power in detecting the mean reversion of stock prices. Perron (1989) argued that if there is a structural break, the power to reject a unit root decreases when the stationary alternative is true and the structural break is ignored. Meanwhile, structural changes present in the data generating process, which have been neglected, sway the analysis toward accepting the null hypothesis of a unit root. Therefore, we proceed to test the stock prices by using the SPSM procedure mixed with the Panel KSS unit root test with a Fourier function to investigate the time-series properties of stock prices for the nine transition countries. The SPSM classifies the whole panel into a group of stationary series and a group of non-stationary series. In doing so, we can clearly identify how many and which series in the panel are stationary processes.

Sequential Panel Selection Method

First, a grid-search is performed to find the best frequency, as there is no a priori knowledge concerning the shape of the breaks in the data. We estimate equation [6] for each integer $k = 1, ...5,$ following the recommendations of Enders and Lee (2009) that a single frequency can capture a wide variety of breaks. The residual sum of squares (RSSs) indicates that a single frequency *(k=2)* works best for most of the series (see the fourth column at the Table 3).

Table 3 reports the results of Panel KSS unit root test with a Fourier function on the stock prices where we give a sequence of the Panel KSS statistics with their bootstrap p-values on a reducing panel, the individual minimum KSS statistic, and the stationary series identified by this procedure each time. As we can see from Table 3, the null hypothesis of the unit root in stock price can not be rejected when the Panel KSS unit root test was first applied to the whole panel, producing a value of -1.581 with a very small p-value of 0.307. To check the robustness of our test, we continued the procedure until the last sequence. We found that the Panel KSS statistic all failed to reject the unit root null hypothesis for the rest of sequences. Apparently, the SPSM procedure using the Panel KSS unit root test with a Fourier function provided strong evidence favouring the random walk process in stock prices for the transition countries being studied.

Fig. 1 displays the time paths of the stock price for each country. We can clearly observe structural shifts in the trend of the data. Accordingly, it appears sensible to allow for structural breaks in testing for a unit root (and/or stationary). The

estimated time paths of the time-varying intercepts are also shown in the Fig 1. A further examination of the figures indicates that all Fourier approximations seem reasonable and support the notion of long swings in stock prices.

Economic and Policy Implications

Several important policy implications emerge from our study. First, if the data were erroneously treated as non-stationary and the causality tests for stock prices and other macroeconomic variables were applied to the first difference, then a spurious causality would result. Second, overwhelming evidence in favor of the I(1) non-stationary hypothesis is found, implying that the stock markets in the transition countries (i.e., Bulgaria, the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Romania, and Russia) were characterized by the weak-form EMH. For that reason, it shows the presence of profitable arbitrage opportunities among the stock prices in these transition countries is not possible. Third, our findings suggest that shocks to stock price are permanent. This result implies that following a major structural change in the global financial markets, stock prices will not return to their original equilibrium over a period of time. The fact that stock prices show I(1) non-stationarity indicates that it should not be possible for the series to forecast future movement in stock prices based on past behavior.

Conclusions

Nelson and Plosser (1982) pointed out that whether stock prices are modeled as a trend-stationary or difference-stationary process has important implications vis-à-vis modeling, testing, and forecasting. Studies on this issue are of considerable concern to researchers conducting empirical studies and investors alike. In this empirical study, we employed the SPSM approach, recently proposed by Chortareas and Kapetanios (2009) to assess the time-series properties of stock prices for 9 transition countries during the 2000.10 to 2010.11 period. The combined use of the Panel KSS unit root test with a Fourier function and the SPSM procedure allows us to convey clear conclusions on the stationarity of individual stock price in our study. Empirical results indicate that a unit root in stock prices can not be rejected for all of the transition countries we studied here.

Notes

1. In other words, if stock prices are mean reverting, then short-selling assets that have performed well and buying assets with relatively poor performance in the past (i.e. contrarian strategies) should provide higher returns (see, Debondt and Thaler (1985); Chan (1998); Richards (1997); Balvers and Wu (2006)).
2. In our sample countries, Estonia, Latvia, Lithuania (the three Baltic countries), Poland, the Czech Republic, Hungary (the three Central European countries), Romania, and Bulgaria (Southeastern Europe), and Russia.
3. Enders and Lee (2009) suggest that the frequencies in (6) should be obtained via the minimization of the sum of squared residuals. Their Monte Carlo experiments suggest that no more than one or two frequencies should be used because of the loss of power associated with a larger number of frequencies.
4. Both first generation and second generation panel unit root tests are also employed in our study, however, the results from both the first and second generation panel-based unit root test are mixed. As we stated earlier that panel-based unit root tests are joint tests of a unit root for all members of a panel and that are incapable of determining the mix of I(0) and I(1) series in a panel setting, therefore we do not reports these results here for space constraints, but results are available upon requests.

References

Bai, J. and P. Perron (1998) Estimating and testing linear models with multiple structural changes. *Econometrica, 66,* 47-78.

Bai, J. and P. Perron (2001) Computation and analysis of multiple structural change models, *Journal of Applied Econometrics, 18,* 1-22.

Balvers, R. and Y. Wu (2006) Momentum and mean reversion across national equity markets, *Journal of Empirical Finance, 13,* 24-48.

Becker, R., Enders, W., and Lee, J (2004) A general test for time dependence in parameters, *Journal of Applied Econometrics*, 19, 899-906.

Becker, R., Enders, W, and Lee, J (2006) A stationairy test in the presence of an unknown number of smooth breaks, *Journal of Time Series Analysis*, 27, 3, 381-409.

Burke, S. 1994. Confirmatory data analysis: The joint application of stationarity and unit root tests. University of Reading, 20

Chan, K. C (1988) On the contrarian investment strategy, *Journal of Business, 61,* 147-163.

Cheung, Y. W. and D. Chinn (1996) Deterministic. stochastic and segmented trends in aggregate output: A cross-country analysis, *Oxford Economic Papers. 48,* 134-162.

Debondt, W. and R. Thaler (1985) Does the stock market overreact? *Journal of Finance, 40,* 793-805.

Dickey, D. A. and Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, 49: 1057-1072.

Enders, W. and Lee, J (2004) Testing for a unit root with a nonlinear Fourier function, working paper, Department of Economics, Finance & Legal Studies, University of Alabama, Tuscaloosa, AL, USA.

Enders, W. and Lee, J (2009) The flexible Fourier form and testing for unit roots: An example of the term structure of interest rates, working paper, Department of Economics, Finance & Legal Studies, University of Alabama, Tuscaloosa, AL, USA.

Gallant, R. (1981) On the basis in flexible functional form and an essentially unbiased form: the flexible Fourier form. *Journal of Econometrics*, 15:211–353.

Kwiatkowski, D., Phillips, P., Schmidt, P., and Shin, Y (1992), "Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root: How Sure Are We That Economic Time Series Have a Unit Root?" *Journal of Econometrics*, 54, 159-178.

Leybourne, S., Newbold, P and Vougas, D. (1998) Unit roots and smooth transitions, *Journal of Time Series Analysis* 19, 8397.

Liu, J., S. Wu, and J. V. Zidek (1997) On segmented multivariate regression, *Statistic Sinica. 7,* 497-525.

Ludlow, J., and Enders, W. (2000) Estimating non-linear ARMA models using Fourier coefficients. *International Journal of Forecasting*, 16, 333–347.

Maddala, G. S. and S. Wu (1999) A comparative study of unit root tests with panel data and a new simple test, *Oxford Bulletin of Economics and Statistics. Special Issue,* 631-652.

Narayan, P. K. (2008) Do shocks to G7 stock prices have a permanent effect? Evidence from panel unit root tests with structural break, *Mathematics and Computers in Simulation. 77,* 369-373.

Narayan, P. K. and R. Smyth (2005) Are OECD stock prices characterized by a random walk? Evidence from sequential trend break and panel data models, *Applied Financial Economics, 15,* 547-556.

Nelson, C. and C. Plosser (1982) Trends and random walks in macroeconomic time Series, *Journal of Monetary Economics, 10,* 139-162.

Newey, W. and K. West (1987) A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica, 55,* 703-708.

O'Connell, P. G. J. (1998) The overvaluation of purchasing power parity, *Journal of International Economics, 44,* 1-20.

Pascalau, R. (2010) Unit root tests with smooth breaks: an application to the Nelson-Plosser data set, *Applied Economics Letters*, 17, 565-570.

Perron, P. (1989) The great crash, the oil price shock and the unit root hypothesis, *Econometrica, 57,* 1361-1401.

Perron, P and T. J. Vogelsang (1992) Testing for a unit root in a time series with a changing mean: Corrections and extensions, *Journal of Business Economics and Statistics, 10,* 467-469.

Richards, A. J (1997) Winner-loser reversals in national stock market indices: Can they be explained? *Journal of Finance, 52,* 2129-2144.

Ucar, N. and Omay, T. (2009) Testing for unit root in nonlinear heterogeneous panels, *Economics Letters,* 104, 5-8.

Table 1: Summary statistic: $\Delta \ln P$ (2000.10-2010.11)

Statistic	Bulgaria	Czech	Estonia	Hungary	Latvia	Lithuania	Poland	Romania	Russia
Mean	-0.009	-0.012	-0.009	-0.012	-0.010	-0.009	-0.007	-0.012	-0.010
Median	0.002	0.005	0.003	0.003	0.002	0.003	0.004	0.004	0.007
Max.	0.294	0.156	0.160	0.149	0.242	0.248	0.229	0.157	0.342
Min.	-5.832	-7.062	-6.530	-7.003	-5.966	-5.986	-4.809	-8.571	-7.379
Std. Dev.	0.258	0.310	0.287	0.308	0.263	0.263	0.215	0.377	0.327
Skew.	-21.938	-22.430	-22.430	-22.372	-22.188	-22.384	-21.256	-22.411	-21.928
Kurt.	495.528	510.281	510.352	508.558	503.023	508.979	474.987	509.732	495.198
J-B	5348.62***	5673.2***	5674.78***	5634.8***	5512.3***	5644.1***	4912.6***	5660.9***	5341.4***

Note: *** and ** indicate significance at the 0.01 and 0.05 levels, respectively, $\Delta \ln P = \ln P_t - \ln P_{t-1}$.

Table 2: Univariate unit root tests : (2000.10-2010.11)

			Level			*First Differences*
Country	***ADF***	***PP***	***KPSS***	***ADF***	***PP***	***KPSS***
Bulgaria	*-1.5265*	*-1.6352*	*1.6462****	*-9.6558****	*-22.0634****	*0.6161***
	(3)	*(12)*	*[17]*	*(2)*	*(11)*	*[12]*
Czech	*-1.2878*	*-1.1408*	*2.0730****	*-10.2693****	*-21.3509****	*0.1993*
	(4)	*(8)*	*[17]*	*(3)*	*(7)*	*[8]*
Estonia	*-1.2829*	*-1.3493*	*1.8647****	*-12.5998****	*-19.3451****	*0.2218*
	(2)	*(9)*	*[17]*	*(1)*	*(8)*	*[9]*
Hungary	*-1.0645*	*-1.1374*	*1.7895****	*-21.7624****	*-21.7956****	*0.1212*
	(0)	*(8)*	*[17]*	*(0)*	*(8)*	*[8]*
Latvia	*-1.7155*	*-1.7405*	*1.4042****	*-7.2832****	*-21.1564****	*0.3388*
	(8)	*(7)*	*[17]*	*(7)*	*(9)*	*[7]*
Lithuania	*-1.2427*	*-1.1806*	*1.8963****	*-8.1930****	*-20.5466****	*0.1779*
	(3)	*(13)*	*[17]*	*(2)*	*(12)*	*[13]*
Poland	*-1.2261*	*-1.2779*	*2.1310****	*-24.7027****	*-24.6538****	*0.0988*
	(0)	*(8)*	*[17]*	*(0)*	*(8)*	*[7]*
Romania	*-2.3431*	*-2.1423*	*2.0554****	*-10.5771****	*-22.9353****	*0.4949***
	(0)	*(9)*	*[17]*	*(2)*	*(10)*	*[10]*
Russia	*-1.4416*	*-1.4800*	*2.3317****	*-21.6767****	*-21.8701****	*0.1308*
	(0)	*(9)*	*[17]*	*(0)*	*(8)*	*[9]*

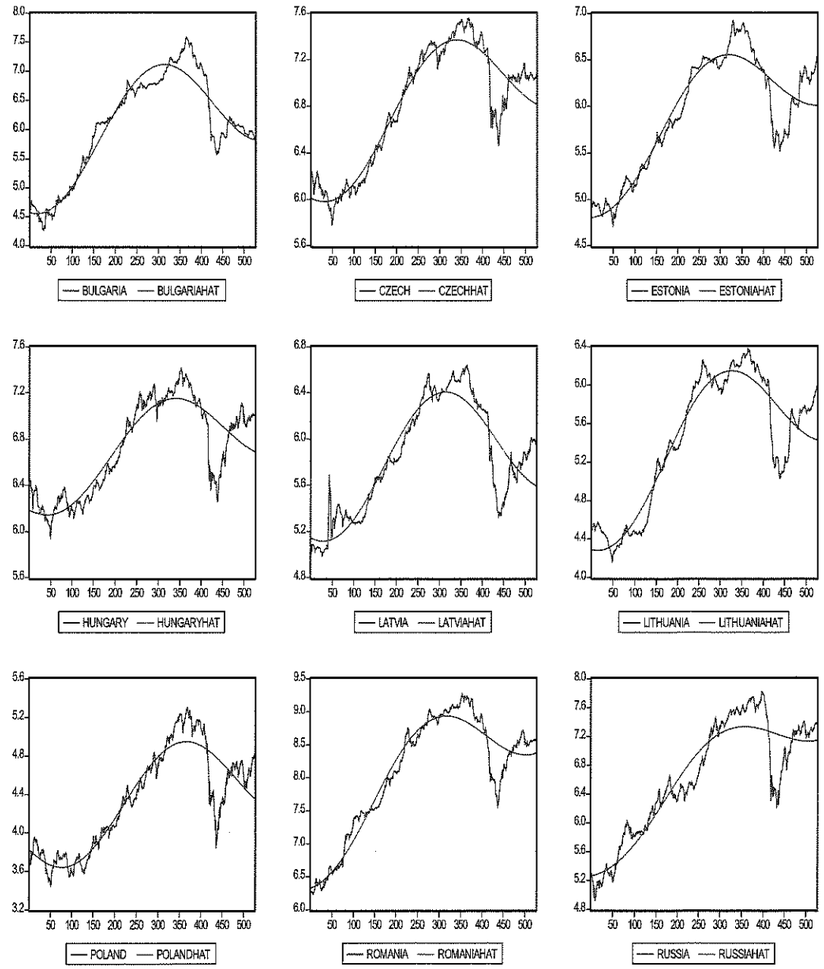
*Note: *** and ** indicate significance at the 0.01 and 0.05 levels, respectively. The number in parenthesis indicates the lag order selected based on the recursive t-statistic, as suggested by Perron (1989). The number in the brackets indicates the truncation for the Bartlett Kernel, as suggested by the Newey-West test (1987).*

Table 3: Results of KSS Test with Fourier function on Stock Prices

Sequence	OU statistic	Min. KSS	Fourier(K)	Series
1	-1.581(0.307)	-2.722	2	Romania
2	-1.439(0.456)	-1.659	2	Lithuania
3	-1.407(0.473)	-1.628	2	Bulgaria
4	-1.370(0.422)	-1.618	2	Poland
5	-1.321(0.470)	-1.616	2	Estonia
6	-1.247(0.505)	-1.490	2	Latvia
7	-1.165(0.502)	-1.403	2	Russia
8	-1.047(0.597)	-1.248	2	Hungary
9	-0.846(0.623)	-0.846	2	Czech

Note: Entry in parenthesis stands for the asymptotic p-value. The significance level is 10%. The maximum lag is set to be 8. The asymptotic p-values are computed by means of Bootstrap simulations using 5000 replications. Fourier (k) is chosen by minimum sum square of residual for Fourier function.

Figure 1: Logs of Stock Price Indices and Fitted Nonlinearities for the Nine Transition Countries



Market Timing Technique: A Trader's Mantra in Volatile Market?

William Cheng, Troy University

Abstract

Over the years, technical analysis has been viewed by market professionals and academicians as market astrology or a kind of voodoo. However, the scary market rumbles did more than just reduce the size of stock portfolio; they also challenged the bases of conventional market thinking. Using data and analytical tools from Bloomberg, the paper examines and compares the holding period return of the Directional Movement Indicator (DMI) -- a popular market timing technical strategy with a conventional buy and hold rule in U.S. stock market.

Introduction

There are two basic investment philosophies, ignoring the efficient market theory, underlying all traders= market strategies: fundamentalism and technicalism. Fundamental analysis contends that security price is determined by fundamental economic factors such as earnings, dividends and sales. Fundamentalists study the cause of market and ask why. Technical analysis contends that past price movements will indicate future price movements. Technicians study the effect of the market and do not care why. Martin J. Pring, a well-respected technical analyst, defines technical analysis as follows:

> *The technical approach to investment is essentially a reflection of the idea that prices move in trends which are determined by the changing attitudes of investors toward a variety of economic, monetary, political, and psychological forces. The art of technical analysis-for it is an art-is to identify trend changes at an early stage and to maintain an investment posture until the weight of the evidence indicates that the trend has reversed..*(Pring, 1990)

Technical Analysis is the study of historical price action, primarily through the use of charts, for the purpose of forecasting future price direction in order to guide the trading. There are three major premises to technical analysis: first, market action discounts everything. Anything that can affect the price---whether fundamental or political news, psychological perceptions, fads or fashions, is actually through the interaction of demand and supply and reflected into the price of the security. Therefore, a study of price action is all that is needed. Second, prices move in trends. The main purpose of charting price action is to identify trends, usually for purpose of trading in the direction of those trends. Third, history repeats itself.

Chart patterns reflect certain pictures that appear on price charts. The pictures reveal the bullish or bearish psychology of the market. Since the patterns have worked in the past, it is assumed that they will continue to work in the future. Another way of saying this that the key to understanding the future lies in the study of the past, or, that the future is just a repetition of the past.(Bensignor, 1998.)

Over the years, technical analysis has been viewed by market professionals and academicians as a market astrology or a kind of voodoo. However, the scary market rumbles of late last summer did more than just reduce the size of stock portfolio, they also challenged the bases of conventional market thinking.(Haugen, 1999) Using data and analytical tools from Bloomberg, the paper examines and compares the holding period return of a popular market timing technical strategy with a conventional buy and hold rule in 1998's U.S. stock market.

Data and Methodology

In the study, we chose the well-known Standard and Poor's 500 Index price movements to test and compare the holding period returns generated by a technical indicator and a common sense buy and hold strategy. All data and analysis are obtained and done through Bloomberg Financial Service. Bloomberg provides 24-hour, instant and current financial, economic and political information covering markets around the globe. It also provides analytics, historical data, up-to-minute news reports, economic statistics and political commentaries. Constant upgrades and enhancements of the system are some of the most valuable attributes of the Bloomberg service.

It shows that the market rallied up to mid-July and then followed by a severe correction downward. In response, Federal Reserve cut federal funds rate for three consecutive times from 5.5% to 4.75%. The market rebounded back into

positive side and went upward all the way through the fourth quarter. This ups and downs also revealed in table 1. which reports the quarterly returns of S&P 500 Index in 1998.

Table 1: Quarterly Returns of the S&P500 Index in

Quarter	Return
Q1	13.53%
Q2	2.91%
Q3	-10.30%
Q4	20.87%

Source: Bloomberg, 1998

The technical indicator used in the analysis is the Directional Movement Indicator (DMI).The Directional Movement Indicator, developed by Welles Wilder, is based upon the assumption that when the trend is up, today's high price should be higher than yesterday's high price. Conversely, when the trend is down, today's low price should be lower than yesterday's low price. It assists in determining if a trend is progressing and also quantifies the strength of the trend so users can identify potential opportunities or alter existing strategies.

Specifically, we first calculate +DI, the "positive" directional movement indicator which is the difference between today's high price and yesterday's high price. Then we estimate DI, the "negative" directional movement indicator is the difference between today's low price and yesterday's low price. A buy indication may be suggested when the +DI crosses up through the BDI. A sell indication may be suggested when the BDI crosses up through the +DI.

Then, we use these indicators to form buy and sell moves through 1998. The result is reported in the next section. In addition, the greater the difference between +DI and -DI, the more the security is trending. This can be done by calculate the Average Directional Movement Index, (ADX), which is derived directly from +DI and BDI, and measures the extent to which a security is trending. Therefore, the higher the ADX, the more trending the security is experiencing (whether up or down). Thus the ADX does not distinguish between a rising or falling trend; it measures the strength of trend ---regardless of direction. It is perfectly normal for the ADX to be rising sharply when the stock's price is failing, because by rising it is indicating the increasing strength of the downtrend.

Empirical Results: Technical vs. Conventional

Table 2 shows the buy and sell actions and dates of entry and exit for technical analysis, buy and hold strategy and under perfect information. Table 3 compares the holding period return of the three approaches. Based on DMI technical trading rule, there were three signals identified in 1998 to either enter or exit the market. The holding period return generated by following this technical analysis is 27.16%. However, if a trader simply chose to buy and hold through 1998, then he/she would gain a total 26.07% return. In addition, if we assume under perfect information or hindsight case, a trader would yield a total of 57.75% return in 1998.

Table 2: Dates and Actions of Various Strategies

Technical Analysis(DMI)		Buy and Hold		Perfect Information	
Date	Action	Date	Action	Date	Action
2/2/98	Buy	1/2/98	Buy	1/9/98	Buy
5/18/98	Sell	12/31/98	Sell	7/17/98	Sell
6/25/98	Buy			8/31/98	Buy
7/24/98	Sell			12/29/98	Sell
10/19/98	Buy				
12/31/98	Sell				

Source: Bloomberg, 1998

Table 3: Comparison of returns of Various Strategies

Technical Analysis (DMI)	Buy and Hold	Perfect Information
27.16%	26.07%	57.75%

Source: Bloomberg, 1998.

Conclusion

We have found that traders who chose one of the technical measures, the directional movement indicator, to find clues regarding the future price direction of the Standard and Poor's 500 Index in 1998, would not yield higher returns than conventional approach. The holding period return of technical indicator (27.16%) is just a little more than return generated through buy and hold strategy (26.07%). It could be even worse if the transaction cost (commissions) were added.

However, the directional movement indicator did provide sell signal during the third quarter of 1998, when the market price were plummeting drastically. In other words, technical analysis is more flexible and able to react to market movement and changing patterns.

The case is still inconclusive. We only selected and tested one technical indicator, namely directional movement indicator, exclude numerous other widely used indicators. Only the U.S. equity market in 1998 was used and a single market index, Standard and Poor's 500 Index was employed in the analysis to compare holding returns from technical strategies with buy and hold investment strategy.

In addition, technical analysis is very subjective. A lot of it comes down to just being intuitive. In our study, we only apply DMI's buy and sell recommendations mechanically without questioning, combining and searching many other market clues of trends or momentum. All can be added and improved in the future research.

References

Bensignor, Rick. "An Introduction to Technical Analysis", *Bloomberg Financial Markets*, Feb. 17, 1998.
Pring, Martin J. *Technical Analysis Explained.* (New York McGraw-Hill, 1991), pp. 2-3.
Fama, Eugene F., and Kenneth R. French. "Permanent and Temporary Components of Stock Prices." *Journal of Political Economy* 98 (1988): pp. 247-273.
Haugen, Robert A. *The Inefficient Stock Market.* 2nd ed.(Prentice-Hall, 2001).
Lehmann, Bruce N. "Fads, Martingales, and Market Efficiency." *Quarterly Journal of Economics*, Feb. 1990, pp 1-28.

The Effects of Events Connected with the Greek Debt Crisis on Greek Credit Default Swap Prices

Jordan G. Kratter, Quinnipiac University

Abstract

In 2011, the price of credit-default swaps in Greece has risen due to fears that the country will default on its debt. The delayed acceptance of various austerity measures and problems with a European bailout have added to the volatility of these securities. This paper investigates the relationship between the market risk premia on these securities and public policy changes in Greece. Specifically, this is an event study that will look at fluctuations in the price of Greek credit-default swaps in relation to the process of reviewing and accepting announced austerity measures and the riots that have resulted from these events. To control for risk factors that are not unique to Greece, this study analyzes the spread between price of Greek credit-default swaps and the price of credit-default swaps in a benchmark nation, Germany.

This paper employs the highly regarded event study methodology developed by Fama, Fischer, Jenson, and Roll (1969), as well as extensions that have been made to this technique since its inception. This method will be used to recognize normal price changes in these swaps, which in turn can be used to identify abnormal price changes. After identifying these abnormal changes, this paper will use econometric techniques to measure their statistical significance. This paper will test my hypothesis that the riots in Greece have had a negative effect on the price of credit-default swaps, and will shed light on the quantitative significance of such events.

Introduction:

Credit derivatives are widely used, and relatively new, financial instruments that allow companies to manage credit risk by trading securities. The most common credit derivative, the credit default swap (CDS), transfers the loan's default risk to the seller, who guarantees the buyer's loan. If the loan defaults, the seller will compensate the buyer with the face value of the loan and takes possession of the loan, an exchange referred to as the swap. The "price" of a CDS is actually the rate that the buyer must pay to the seller and is referred to as the CDS "spread," which is measured in basis points, where one basis point is equal to 0.01% (Skinner, 2005).

For example, suppose the CDS spread on ABC Corporation's debt is 500 basis points and a hedge fund purchases ABC's bonds with a face value of $100 million. To protect the investment, the hedge fund buys CDSs from XYZ Bank at a cost of $5 million per year. Then if ABC Corporation were to default on its debt and could only pay the hedge fund $60 million of the bond's face value, XYZ Bank must provide the hedge fund with $100 million in exchange for the securities now worth $60 million. This effectively restores to the hedge fund the $40 million it would have otherwise lost. This example is quite simple, but CDSs can be used to insure bonds, mortgages, bundles of mortgages and more. For this study, only a basic theoretical understanding of CDSs is necessary.

It should also be noted that because CDSs are not technically a form of insurance, they are not highly regulated. In particular, the buyer does not need to own the underlying security and the seller does not need to hold reserves to pay off the buyers (Grabowski, 2008). Yet, the practical purpose of a CDS is to "insure" some underlying security. Therefore this study will refer to CDSs as a form of insurance, although technically they are not.

CDSs, introduced in the 1990's, became very popular after about a decade, and by 2007, the notional value of all CDSs was over $62 trillion (ISDA Market Survey, 2009). By this time, the housing bubble had burst, mortgages were defaulting, and CDS sellers were obligated to pay out trillions, but did not always have the funds to do so. A recession spread across the world, and within three years that notional amount had dropped by more than half (Summary of Recent Survey Results, 2010).

At the conclusion of World War II, the Greek economy was in shambles. Greece began rebuilding, taking on tremendous amounts of debt. Over the years, Greece's borrowing increased. By 1995, the government debt was equal to 73.52% of GDP, and by 2010, it had risen to 142.8% of GDP (Paris, Dedes, and Lampridis, 2011). Because the country has so much debt, its probability of default has increased significantly. As a result, many investors who did not own Greek debt began buying CDSs on the country's debt, entitling them to be paid the face value of the government's bonds in the case of a default. Greece currently accounts for $70.8 billion of the outstanding CDSs (Factbox: What Happens If Greece Defaults?, 2012). As a result, CDS prices have skyrocketed, and in December 2011, a financial phenomenon occurred in which insurance on Greek debt briefly became more expensive than the debt itself.

This paper analyzes the impact of several types of political and economic events on Greek CDS prices, employing the highly regarded event study methodology originally developed by Fama, Fischer, Jenson, and Roll (FFJR) in 1969, as well as extensions that have been made to this technique since its inception. This method will be used to identify normal price changes in these CDS spreads, by comparison to a benchmark rate. Then, price changes specific to Greek CDSs will be isolated. The abnormal returns, unique to Greece, represent changes in the risk premiums associated with these securities, most likely in response to the events studied. In addition, this paper will analyze the statistical significance of these abnormal price changes and draw conclusions from the results.

This paper will use the German CDS spread as the benchmark rate. Since Standard & Poor's gives Germany a credit rating of AAA (Germany (Federal Republic Of) (Unsolicited Ratings), 2012), which means it has a very low chance of defaulting on its debt, its CDS spread can be seen as the European Union's (EU) benchmark CDS rate. As a result, the Greek CDS spread can be compared to the German CDS spread to find the risk premium associated with Greek debt. When an event occurs that is specific to Greece, the movement of the Greek CDS spread relative to the German CDS spread will represent that event's impact on the default risk of Greek debt.

The event study will cover the daily 5-year Greek and the daily 5-year German CDS spreads. Over this time period, based on the numbers in Table 1, the daily 5-year German CDS spread increased by nearly four times its original rate, while the daily 5-year Greek CDS spread increased by nearly 30 times its original rate. This paper will study events that may have contributed to the enormous increase in Greece's CDS risk premium. These events have been divided into four groups: strikes and riots, credit rating changes, the convening or disbanding of meetings for austerity measures and aid, and the announcement of austerity measures and aid from outside sources. From the analyses, conclusions will be drawn as to whether or not specific events had impacts on the CDS spreads, how large those impacts were, and what types of events had the largest impacts. In addition, the study will compare the effects of positive and negative events to see if one has a greater impact on CDS spreads than the other.

Another important event that may take place in the near future needs to be mentioned. Private investors were asked to take a voluntary 50% "haircut" on Greek debt. Currently, a deal is being worked out that could result in a voluntary restructuring of Greek debt, which would not technically be considered a default. As a result, CDS contracts would not be terminated, and the face value of the debt would not be paid back, altering the definition of a default. This event would lower the conceptual value of a Greek CDS and would cause a dramatic drop in the Greek CDS spread. However, if a deal is not reached, the Greek government would be forced to default on its debt, causing the swap to occur. The outcome could significantly change all CDS pricing.

Literature Review:

The event study methodology originated with FFJR (Binder, 1998). Since then, there have been many additions and revisions to the method. For this paper, the introduction of the dummy variable to the FFJR model is the most relevant of these contributions (Izan, 1978).

FFJR used the market model equation (Binder, 1998):

$$R_{it} = \alpha_i + \beta_i R_{mt} + u_{it}$$

to analyze the abnormal returns of a stock. Here the error variable, u_{it}, is used as an estimator for the abnormal returns. In essence, the abnormal returns are modeled as prediction errors from the market model equation.

Instead of looking at abnormal returns as prediction errors, a dummy variable can be added to the equation, which then includes the event period as part of the sample period. The following equation results:

$$R_{it} = \alpha_i + \beta_i R_{mt} + \gamma_i D_t + u_{it}\,.$$

This equation uses the coefficient, γ_i, to represent the abnormal returns for a security in the case of an event, while D_t represents whether or not that event occurs. Izan (1978) created this model to examine a portfolio of firms, all of which experienced some events, by setting the dependent variable equal to the equally weighted portfolio's return.

This equation has been altered for purposes of this paper to the following:

$$P_{GrCDSPt} = \alpha + \beta P_{GerCDSPt} + \gamma_1 D_1 + u_t,$$
$$P_{GrCDSPt} = \alpha + \beta P_{GerCDSPt} + \gamma_1 D_1 + \gamma_2 D_2 + \gamma_3 D_3 + {}_4 D_4 + u_t$$

and

$$P_{GrCDSPt} = \alpha + \beta P_{GerCDSPt} + \gamma_1 D_1 + \gamma_2 D_2 + u_t\,.$$

These formulas will be further explained in the methodology section of this paper.

Greatrex (2008) conducted a similar event study in which she analyzed the informational efficiency of the CDS market by comparing the CDS market's response to earnings announcements to the stock market's response. She also compared the response of both markets to S&P credit rating changes. Although Greatrex's (2008) study is similar to the one in this paper in that it is an event study that examines fluctuations in the CDS market, there are key differences. First and foremost, Greatrex (2008) analyzed the effect that events had on price changes in a cross-section of firm's CDS spread, while this paper looks at the Greek CDS market. Also, Greatrex (2008) compared CDS spread movement to stock market movement, while this study will compare the movement of Greek CDS spreads to German CDS spreads, using a different mathematical model. This results in a disparity in the methodologies used.

It remains to be seen in the results section if this study's conclusions are parallel to those of Greatrex (2008). She concluded that events had a statistically significant impact on the price fluctuations of these securities. In addition, she found that the market reacted more strongly to negative news than to positive news and that larger events had greater affects on price movement. She concluded that the CDS market reacts more efficiently to credit rating events than to earnings announcements.

Hull, Predescu, and White (2004) examined the relationship between CDS spreads and both bond yields and credit rating announcements. They used the relationship between CDS spreads and bond yields to estimate the 5-year risk-free rate used by participants in the CDS market. Then they compared the risk-free rate to the CDS spreads, which revealed the risk premiums associated with them. They confirmed that there is a relationship between credit rating announcements and abnormal returns, and then found that negative events had a greater statistical impact on CDS spreads than positive events. This is consistent with the findings in Greatrex (2008).

Methodology:

This study analyzes the impact of specified events on the daily 5-year Greek CDS spread in 2011. The events considered have been separated into four groups: strikes and riots, credit rating changes, the convening or disbanding of meetings for austerity measures and aid from outside sources, and the announcement of austerity measures and aid from outside sources. Using the event study methodology and a multivariate regression model, normal price fluctuations on days when no event occurs will be identified and compared to the price changes on days when events do occur. The study will measure the statistical significance of each type of event.

In order to gain perspective on the significance of the results, the study will use the daily 5-year German CDS spread as a frame of reference. Although the events studied are specific to Greece, they have an impact on economies around the world. By comparing the Greek CDS spread to the German CDS spread, this study will isolate the impact these events have on the Greek CDS risk premium. The German CDS spread was selected as the benchmark rate because Germany has the highest credit rating and is part of the EU. Standard & Poor's gives Germany a credit rating of AAA, which means there is virtually no chance that it will default on its debt, and investments in Germany's debt have virtually no risk of default. In addition, both countries use the same currency, the Euro. Therefore, this study will consider the German CDS spread to be comprised of only the risk-free rate and general market premiums, and will compare the Greek and German rates to in order to expose the risk premiums specific to Greek debt.

The data being input to the model are the daily 5-year Greek and German CDS spreads. Originally, this study was going to use the 1-year CDS spread, but in the midst of the debt crisis, the 1-year spread exceeded 10,000 basis points because most investors believed Greece would default on its debt in the near future. As a result, the 1-year spreads were no longer traded consistently. Due to their illiquidity, they did not accurately represent the effects of the events studied. As 5-year spreads were actively traded, the study will use them instead.

Data was gathered from 2010 and 2011. In an event study, it is important to have a pre-event data set so that a regression model can properly interpret abnormal price movement (Binder, 1998). In this study, 2010 is being used as the pre-event period. Although the Greek debt crisis began during that year, there were few events as defined by this study. Because the vast majority of the events occurred in 2011, the model will be able to estimate abnormal CDS spread changes reasonably accurately.

The daily 5-year CDS spreads for both Greece and Germany were found in the "CMAN" database. "CMAN" is an external database that is found on "Bloomberg's Financial Database." This data was imported into Microsoft Excel 2011, which was used to perform the multivariate regression. The textbooks, Essentials of Econometrics by Damodar N. Gujarati and Pricing and Hedging Interest & Credit Risk Sensitive Instruments by Frank Skinner, were used for the research on CDS and econometric theory. ProQuest's ABI Inform was used as a research tool. First, ABI Inform's collection of newspaper

subscriptions was used to identify dates in which events took place. Second, its journal subscriptions were used to find studies that have already been done on the effects that events have on CDS spreads.

This analysis uses the event study methodology originated by FFJR, which has been modified several times, most importantly, for purposes of this study, by the addition of the dummy variable by Izan (1978). The dummy variable is a zero-one variable that is zero when an event does not take place and one when it does. The first model used in the study is similar to the model developed by Izan (1978):

$$P_{GrCDSPt} = \alpha + \beta P_{GerCDSPt} + \gamma_1 D_1 + u_t . \quad (1)$$

This model will reveal whether there are in fact abnormal fluctuations in the Greek CDS spread on days when an event occurs. If there are, as is hypothesized, the study will break down the events into categories and examine the impacts of various types of events. The model expresses the Greek CDS spread as a function of the German CDS and a dummy variable. α represents the intercept of the model, which has little theoretical value because the German CDS spread would never be equal to zero. β represents the risk premium with respect to German debt that an investor must pay to insure Greek debt, without taking into account abnormal returns from events. D_1 represents the dummy variable that indicates when an event occurs, and γ_1 represents the average abnormal returns on Greek CDS when an event occurs.

Because the previous model provides results in basis points, it will provide answers that can only be interpreted in relation to the CDS spread. However, the results could be more easily interpreted if they were in percentage form. In addition, the previous model is most accurate for linear relationships, which will not necessarily the case here. Therefore, the following model:

$$\ln(P_{GrCDSPt}) = \alpha + \beta \ln(P_{GerCDSPt}) + \gamma_1 D_1 + u_t. \quad (2)$$

will also be used. By taking a natural logarithm of the prices, this model will be better able to account for non-linear relationships. Also, this model will express the impact of the events, γ_1, as a percentage change.

The third model used to understand the relationship between the occurrence of an event and CDS spread movement is:

$$(P_{GrCDSPt} - P_{GerCDSPt}) / P_{GerCDSPt} = \alpha + \beta P_{GerCDSP} + \gamma_1 D_1 + u_t. \quad (3)$$

Instead of measuring the basis point spread, this model identifies the Greek CDS *risk premium* in relation to the German CDS spread, determining the average percentage change in the in Greek CDS risk premium with respect to the German CDS. This is an important distinction because it isolates the risk premium, which is the true focus of this study.

The fourth model developed for this study is:

$$P_{GrCDSPt} = \alpha + \beta P_{GerCDSPt} + \gamma_2 D_2 + \gamma_3 D_3 + \gamma_4 D_4 + \gamma_5 D_5 + u_t. \quad (4)$$

This model expresses the Greek CDS spread as a function of the German CDS and four dummy variables. The data collected provides all of the German and Greek daily 5-year CDS spreads, while the model will solve for the remaining variables. β represents the risk premium with respect to German debt that an investor must pay to insure Greek debt, without taking into account abnormal returns from events. D_2 represents the dummy variable for strikes and protests, and γ_2 represents the average abnormal returns on Greek CDS when strikes and protest occur. D_3 represents the dummy variable for announcements regarding credit rating changes, and γ_3 represents the average abnormal returns on Greek CDS when Greece's credit rating has been adjusted. D_4 represents the dummy variable for convening or disbanding meetings concerning austerity measures and aid from outside sources, and γ_4 represents the average abnormal returns on Greek CDS when these meeting occur or are unexpectedly ended. D_5 represents the dummy variable for the announcement of austerity measures or aid from outside sources, and γ_5 represents the average abnormal returns on Greek CDS when these announcements occur. Finally, u_t represents the error term, a normal piece of a regression model, as a regression is simply an average.

This model uses five independent variables, one for each of the four event categories and one for the German CDS spread, and one dependent variable, the Greek CDS spread. A model such as this could fall into the "dummy variable trap," which occurs when more than one dummy variable is equal to one at the same time (Gujarati and Porter, 2010). If this happens, there will be an error in the results. For this study, the model would only fall into the dummy variable trap if more than one type of event on the same day. This only happens on one event day out of the 45 event days identified in this study. Therefore, while on that day there will be some errors in the results, the errors will be minimal and the validity lost will be

negligible, particularly because conclusions will be drawn from having a general idea of each event's effect, as opposed to its exact mathematical impact.

Similar to the modifications to the first model, this model will be changed twice to obtain a better understanding of the impact each type of event studied had on the Greek CDS spread. The fifth model:

$$\ln(P_{GrCDSPt}) = \alpha + \beta \ln(P_{GerCDSPt}) + \gamma_2 D_2 + \gamma_3 D_3 + \gamma_4 D_4 + \gamma_5 D_5 + u_t, \quad (5)$$

uses natural logarithms to determine the average daily percentage change of the Greek CDS spread for all four types of events studied. Again, the use of natural logarithms will better account for non-linear relationships between the variables.

The sixth model, also used to measure the impact of each category of event, is:

$$(P_{GrCDSPt} - P_{GerCDSPt}) / P_{GerCDSPt} = \alpha + \beta P_{GerCDSPt} + \gamma_2 D_2 + \gamma_3 D_3 + \gamma_4 D_4 + \gamma_5 D_5 + u_t . \quad (6)$$

This model identifies the Greek risk premium, in relation to German CDS spread, determining the average percentage change of the Greek CDS risk premium with respect to the German CDS.

The seventh model used in the study is:

$$P_{GrCDSPt} = \alpha + \beta P_{GerCDSPt} + \gamma_6 D_6 + \gamma_7 D_7 + u_t . \quad (7)$$

This model will be used to identify any abnormal returns associated with positive and negative events. In this model, D_6 is the dummy variable for the occurrence of a positive event and D_7 is the dummy variable for the occurrence of a negative event. The rest of the variables are consistent with the fourth model. Any abnormal returns, γ_6 and γ_7, can be compared to determine whether positive or negative events had a larger impact on Greek CDS spread.

This model was also modified twice to better understand the impacts of positive and negative effects. The eighth model is shown by the following formula:

$$\ln(P_{GrCDSPt}) = \alpha + \beta \ln(P_{GerCDSPt}) + \gamma_6 D_6 + \gamma_7 D_7 + u_t . \quad (8)$$

Again, natural logarithms are used to determine the average daily percentage change in the Greek CDS spread, for both positive and negative events. Also, the use of natural logarithms will better account for non-linear relationships between the two variables.

The final model that will be used is:

$$(P_{GrCDSPt} - P_{GerCDSPt}) / P_{GerCDSPt} = \alpha + \beta P_{GerCDSPt} + \gamma_6 D_6 + \gamma_7 D_7 + u_t . \quad (9)$$

This model identifies the Greek risk premium, in relation to German CDS spread, determining the average percentage change of the Greek CDS risk premium with respect to the German CDS.

The statistical significance of a coefficient is determined by its respective p-value. Typically, a p-value less than 0.05 or 0.1 is accepted as being statistically significant. This study will accept p-values below 0.05, although it will still consider p-values less than 0.1. The study will also look at the value of r^2, although it is not an absolute measure of linear association. Generally, a high r^2 value shows that the model's estimated values of the dependent are close to the actual values of the dependent variable. On the other hand, a high r^2 does not always indicate that the results are significant, and a low r^2 does not always indicate poor results. While each model's r^2 value will be analyzed, this study's accepted measure of statistical significance will be based on the p-values.

Before conducting the study, a hypothesis was developed for these models. The hypothesis for the first, second, and third models is that the occurrence of an event will have a substantial impact on Greek CDS prices. The hypothesis for the fourth, fifth and sixth models is that the abnormal returns for each category of event will be noticeably different from the other event categories and that the abnormal returns for credit rating announcements would be the greatest because credit rating announcements have an immediate impact on the risk premiums incorporated into the CDS spread, as the spread will have to adjust to incorporate the additional default risk premium. Finally, the hypothesis for the seventh, eighth and ninth models is that negative events will have a greater impact on the CDS spread than positive events, as found by Hull, Predescu, and White (2004) and Greatrex (2008).

Results:

In the review of ABI Inform's newspaper database, 45 events were identified in 2010 and 2011 that fit the description of an event in this study. Of these, three were in 2010, and 42 were in 2011.

The nine regression models produced a variety of results. That is to say, some of the results were highly statistically significant while others were not. The fourth, fifth and sixth model, in which events were broken down into categories, produced the least significant results, although there was still valuable data to be analyzed.

The first three regression models analyzed the impact of an event, of any kind, occurring on Greek CDS spreads. The results, shown in Table 2, are statistically significant for all three models because each coefficient's p-value is below 0.05. The R^2 values are also fairly high, showing that the regression function as a whole fits the data reasonably well. Therefore, the results of the regression can be accepted.

In the first model, the γ_1 term shows that on average, when an event took place, the Greek CDS spread moved an additional 497.68 basis points, compared to days when no event occurred. In relation to the Greek CDS spread, which was between 286.45 basis points and 11,310.285 basis points through 2010 and 2011, as seen in Table 1, this is a significant daily movement. The second model reveals a 37.9% additional change in the Greek CDS spread on days when an event occurs. By any standard, this is an extreme change for a single day. On the other hand, this result helps to explain the overall 2,967% increase in the Greek CDS spread from January 2010 to December 2011. The third model, which found the percentage increase in the Greek CDS risk premium with respect to the German CDS, concluded that the risk premium increased by an additional 9.55% on days that an event occurred. These results demonstrate that the occurrence of one of these events was associated with a substantial change in the CDS spread.

Regression models four, five and six, shown in Table 3, were designed to reveal which category of events had the greatest impact. In these models, not all of the coefficients had p-values low enough to be statistically significant. On the other hand, the R^2 value shows that the overall regression still fit the data reasonably well.

The fourth regression model only shows statistically significant results for strikes and riots. However, this model also shows that out of all of the events, strikes and riots had the largest coefficient, and therefore the greatest effect, on the Greek CDS basis point change. The fifth regression model produced statistically significant results for three types of events: strikes and riots, the convening or disbanding of meetings about austerity measures and outside aid, and announcements regarding austerity measures and outside aid. In addition, although the p-value for the fourth category of event, credit rating announcements, at 0.059, is above this study's accepted value, it is still close to being statistically significant. This model also found that strikes and riots had the most substantial impact on the Greek CDS spread's percentage change, followed by announcements regarding austerity measures and outside aid, and then the convening or disbanding of meetings about austerity measures and outside aid. Credit rating announcements had the smallest effect, one not quite statistically significant. This disproves the hypothesis, that credit rating announcements would have the greatest impact on CDS spreads. The sixth model, which found the risk premium's percentage change for each type of event, only revealed statistically significant results for two types of events: strikes and riots and announcements regarding austerity measures and aid. Once again, this model shows that strikes and riots had the largest impact among all events. Although not all of the findings were statistically significant, it should be noted that the size of the coefficients for each type of event ranked in the same order as in the fifth model.

Table 4 shows the results for models seven, eight and nine, concerning the impact of positive versus negative events. Only models eight and nine produced p-values in this study's accepted range, but the p-values for the seventh model are reasonably close, and the results are worth considering. The R^2 value for these three models shows that the regressions fit the data reasonably well.

The seventh regression shows that the positive events caused a larger basis point change than negative events. This disproves the hypothesis that negative events would have a greater impact and is inconsistent with the findings of Hull, Predescu, and White (2004) and Greatrex (2008), produced in other contexts. Once again, it is important to note that these were not quite statistically significant results. Regression model eight investigated the Greek CDS percentage change in response to positive and negative events, and found that negative events had a greater impact than positive ones. The ninth regression model, which measured the percentage change for the risk premium associated with the Greek CDS spread in relation to the German CDs spread, also found that negative events had a larger impact than positive ones. The conflicting results between these models do not necessarily mean that some results were inaccurate. Because the majority of positive events happened late in the study, after the Greek CDS spread was very high, the greater average change in basis points reflected in the seventh model would not necessarily result in greater average percentage changes.

Conclusions:

This study has examined the effects of specified events on the 5-year Greek CDS spread. Then, this study broke down the events into categories to better understand their impacts. Models one, two and three confirmed that the occurrence of specified events did have a substantial impact on the CDS spreads and added a substantial risk premium to Greek debt. This helps explain why the Greek CDS spread grew by nearly 30 times its value from January 2010 to December 2011, while the German CDS spread grew by only four times its initial spread over the same period. While the Greek debt crisis would have naturally caused the CDS spread to increase by a large amount, some of the events studied have certainly added to the level of uncertainty regarding the chance that the government may default on its debt. The nation's long history of incurring tremendous amounts of debt has resulted in a likelihood of default.

Models four, five and six revealed that strikes and riots had the largest impact on the CDS spread. This is a peculiar result for a few reasons. First, strikes and riots do not have the direct impact on the creditworthiness of debt that other events have. For example, a credit downgrade directly affects the risk premium on debt because it affirmatively declares an increase in its default risk premium. This was the logic behind the hypothesis that credit rating announcements would have the largest impact on the CDS spread. On the other hand, a credit rating downgrade can be anticipated, and it is possible that investors were already factoring in such risk premiums. Second, strikes and riots are incapable of directly changing the debt crisis the way that austerity measures or aid from other countries could. It is surprising that actions being taken to fight the crisis have less of an impact than acts occurring in reaction to the crisis. On the other hand, it can be argued that the extent of strikes and riots has been so great that it had a negative effect on the Greek debt crisis as a whole. When Greek citizens refuse to go to work they are, in fact, reducing national productivity and, consequently, the country's ability to cover its debt. Many strikes and riots are in reaction to tax increases that citizens do not want to pay, and as the national debt must be paid off by the government's tax income, these events could end up worsening the debt crisis. It is ironic that although people are protesting because they are upset with the situation, their actions have made matters substantially worse. In addition, damages created by strikes and riots will force the Greek government to assume expensive repair costs. Finally, strikes and riots can occur suddenly, meaning that they may not be accounted in advance. Although this study did not investigate why each category of events might affect the CDS spread, these results can provide a basis for formulating hypotheses designed to answer these questions.

The seventh model produced interesting results. In the studies done by Hull, Predescu, and White (2004) and Greatrex (2008), negative events had larger impacts on prices of securities than positive ones. However the opposite was true for seventh model. As stated above, this outcome may have resulted because the majority of positive events occurred late in the study, after the Greek CDS spread was very high, and therefore the greater average change in basis points does not necessarily result in a greater average percentage change. Once again, this study did not explore the reasons for such an unusual occurrence. Another possible explanation is that because a CDS spread is a percentage that must be paid to insure some form of debt, there is a limit as to how much a reasonable investor would be willing to pay. In particular, it would be unreasonable to insure a form of debt if the cost of insurance is greater than the debt. For CDS spreads, 10,000 basis points marks the point at which an investor would have to pay the 100% of the amount of debt to insure it. As a result, this 10,000 basis point mark can loosely be viewed as a ceiling for a CDS spread. Amazingly, the spread climbed past this point for several days in December 2011, but then dropped back below this point. It makes sense that negative events would have less of an impact on a CDS spread as the spread approached the 10,000 basis point ceiling. On the other hand, this limit has no affect on the impact of positive events. Although this speculative, it is a hypothesis suggested by the results of this study. It should be noted again that the results from the seventh model were not statistically significant by this study's standards, but were close enough to be considered.

The eighth and ninth models produced statistically significant results that were more consistent with the findings of Predescu, and White (2004) and Greatrex (2008). Both of these models found that abnormal percentage changes for the Greek CDS spread were greater for negative events than positive events. This is more relevant than the results from the seventh model because the percentage change shows the abnormal return, observed from an event studied, with respect to that day's spread. These results were statistically significant.

The Greek debt crisis has been at the center of attention of the financial world throughout 2010 and 2011. Despite the efforts of the Greek government, as well as aid from many other countries, Greece may well be forced to default on its debt. Many events, particularly strikes and riots, have exacerbated this problem. Surely, without any of the events studied, the debt crisis would still be out of control, but the cost of insuring this debt would not be as high. In addition, a CDS spread can be used as a tool to measure creditworthiness, which means that the events studied have actually lowered the creditworthiness of Greek debt. Unfortunately, Greece is now in shambles, both economically and politically.

References:

Binder, John J. "The Event Study Methodology Since 1969." *The Journal of Quantitative Finance and Accounting* 11 (1998) 111-137.

"Factbox: What Happens If Greece Defaults? | Reuters." *Reuters.com*. Reuters, 19 Jan. 2012. Web. 27 Jan. 2012. <http://www.reuters.com/article/2012/01/19/us-greece-default-idUSTRE80H1NX20120119>.

Garbowski, Mark. "Credit Default Swaps: A Brief Insurance Primer - Insurance - United States." *Mondaq*. Aderson Kill & Olick, P.C., 24 Oct. 2008. Web. 27 Jan. 2012. <http://www.mondaq.com/unitedstates/article.asp?articleid=68548>.

"Germany (Federal Republic Of) (Unsolicited Ratings)." *Standardandpoors.com*. Standard & Poor's, 13 Jan. 2012. Web. 26 Jan. 2012. <http://www.standardandpoors.com/prot/ratings/entityratings/en/us/?entityID=269913>.

Greatrex, C.. *The credit default swap market's determinants, efficiency, and relationship to the stock market.* Diss. Fordham University, 2008. ABI/INFORM Global, ProQuest. Web. 26 Jan. 2012.

Gujarati, Damodar N., Dawn C. Porter. *Essentials of Econometrics*. 4th ed. New York: McGraw Hill/Irwin, 2010. Print.

Hull, John, Mirela Predescu, and Alan White. "The relationship between credit default swap spreads, bond yields, and credit rating announcements. " *Journal of Banking & Finance* 28.11 (2004): 2789. ABI/INFORM Global, ProQuest. Web. 26 Jan. 2012.

"ISDA Market Survey." *ISDA - International Swaps and Derivatives Association, Inc.* Web. 26 Jan. 2012. <http://www.isda.org/statistics/pdf/ISDA-Market-Survey-annual-data.pdf>

Izan, Haji Y. "Mandatory Audit Regulation for Banks: An Empirical Evaluation of Its Effects :I. Introduction. " *The Journal of Business (pre-1986)* 54.4 (1980): 377-396. ABI/INFORM Global, ProQuest. Web. 26 Jan. 2012.

Paris, A., S. Dedes, and N. Lampridis. "Greek Financial Crisis. " *Global Business and Management Research* 3.3/4 (2011): 319-341. ABI/INFORM Global, ProQuest. Web. 26 Jan. 2012.

Skinner, Frank. "Pricing and Hedging Interest and Credit Risk Sensitive Instruments. Burlington: Elsevier Butterworth-Heinemann, 2005. Print.

"Summary of Recent Survey Results." *ISDA - International Swaps and Derivatives Association, Inc.* Web. 26 Jan. 2012. <http://www.isda.org/statistics/recent.html>.

Appendix:

Table 1: Summary of Statistics

Data Inputs	*Mean*	*Std. Dev*	*Minimum*	*Maximum*
Greek CDS (bp)	1900.642	2353.06	251.320	11,310.285
German CDS (bp)	53.356	22.17	24.500	121.510
LN(Greek CDS) (%)	7.059	0.91	5.526	9.333
LN(German CDS) (%)	3.9040.37	0.37	3.197	4.799
Greek CDS_{rp} (%)	27.826	21.26	6.456	107.388

* table displays a statistical summary of the raw data
* prices for Greek and German CDSs are represented in basis points

Table 2: Results for Test 1

	Model 1		*Model 2*		*Model 3*	
Variable	*Coefficient*	*T stat*	*Coefficient*	*T stat*	*Coefficient*	*T stat*
Intercept	-2999.86	-22.73*	-1.16	-5.72*	-10.99	-7.19*
German Price	91.06	39.04*	2.10	40.28*	0.712	26.35*
Dummy 1	497.68	2.68*	0.38	5.49*	9.55	4.43*
R^2	0.762		0.779		0.608	

* significant at the 1% level , ** significant at the 5% level, *** significant at the 10% level

Table 3: Results for Test 2

	Model 1		*Model 2*		*Model 3*	
Variable	*Coefficient*	*T stat*	*Coefficient*	*T stat*	*Coefficient*	*T stat*
Intercept	-3006.71	-22.52*	-1.18	-5.77*	-11.21	-7.25*
German Price	91.21	38.65*	2.10	40.03*	0.72	26.20*
Dummy 2	667.50	1.90***	0.44	3.40*	12.25	3.00*
Dummy 3	240.73	0.58	0.29	1.89***	7.44	1.55
Dummy 4	503.59	1.26	0.31	2.12**	7.75	1.68***
Dummy 5	364.32	1.37	0.36	3.68*	8.13	2.63*
R^2	0.762		0.780		0. 609	

* significant at the 1% level , ** significant at the 5% level, *** significant at the 10% level

Table 4: Results for Test 3

	Model 1		*Model 2*		*Model 3*	
Variable	*Coefficient*	*T stat*	*Coefficient*	*T stat*	*Coefficient*	*T stat*
Intercept	-3002.62	-22.54*	-1.18	-5.75*	-11.14	-7.21*
German Price	91.14	38.65*	2.10	39.90*	0.72	26.16*
Dummy 6	447.46	1.84***	0.33	3.46*	8.32	2.77*
Dummy 7	421.08	1.78***	0.39	4.48*	9.67	3.52*
R^2	0.762		0.778		0.608	

* significant at the 1% level , ** significant at the 5% level, *** significant at the 10% level

Towards an Intuitive Explanation of the Black-Scholes Volatility Model

Richard Lewin, Rollins College

Abstract

Despite rigorous definitions available in many advanced texts, undergraduate students repeatedly struggle to explain how the Black-Scholes option-pricing model works. Students questions are often centered around 'but how does it actually work', or 'how can I explain it at an interview in a way that shows that I actually understand it?' To respond with the proper technical derivation may not be the best way of meeting their needs, at least at the undergraduate level where 'math phobia' may well predominate. This paper sets out to provide an intuitive teaching framework to aid student understanding of the historic Black-Scholes option-pricing formula.

Introduction to What the 'Black-Scholes' Option Pricing Model is all About

Essentially, we are trying to find the correct value today of a contract, called an option, whose ultimate value depends on, or is *derived* from, the value of another underlying asset. A popular example being a stock, say the price of Apple Computers stock (ticker symbol APPL, as traded on the NASDAQ) at a future point in time e.g. one year hence. Naturally, the contract may have different values over the next year, depending on whether the stock price of APPL moves up or goes down.

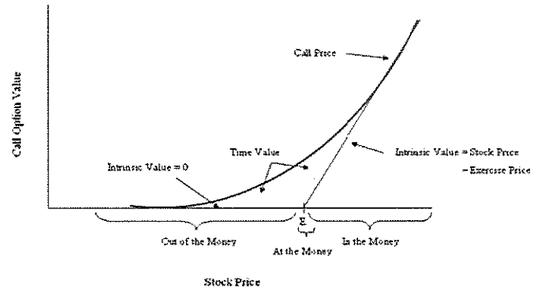
At its simplest level, we can consider that an option provides us with a choice; to do something in the future if it proves to be in our benefit, and to avoid doing something otherwise. In the case of an option, it should be clear that the payoff curve is not symmetrical. The ability to choose only the upside, assuming APPL rises, or more generally to select from amongst only beneficial financial outcomes, is inherently valuable. Not knowing what the likely future price of any asset may be, prior to the expiration of the option contract however, is what makes it difficult to agree upon how much an individual might be prepared to pay someone else, in the form of an insurance (or option premium), in order to allow them to hold such a privileged future position in APPL - while the recipient stands on the opposite side of this opportunity and bears the risk of delivering this asset at exercise or expiry. Intuitively we are going to need some measure, or measures, of the probabilities of the range of the assets price over the contract period. In the case of a stock this is the continuum of values in the future for the stock, after any given period of time as covered under the contract. If we have that mechanism in place, it may be possible to apply an expected present value calculation to get an appropriate price for such a contract.

Celebrated academics Fischer Black and Myron Scholes recognized in early 1973 that although an option price is essentially determined by a few simple factors, (namely the asset's price, the exercise price, the time to expiry, the prevailing interest rate and the volatility of the underlying asset or the measure of the assets riskiness) deriving an option price, could rely on relative pricing, by relating it directly via these variables along with certain simplifying assumptions to the potential range of unanticipated changes in the underlying asset's future price - in our case the future price of APPL stock.

To begin with, we can examine our simple call option as the right but not the obligation to buy an asset at some future date for a known price today. This can typically be presented as a straight line to represent the (risk-free) minimal value of an option at any given time, based on the current prevailing price of the underlying asset (or what we would refer to as the difference or its *intrinsic value;* i.e. mathematically what would the option be worth today if it were to be immediately exercised? We can then draw a wavy line on top of that as an "addition" to this minimum value already shown, to symbolize the potential, yet uncertain, additional value of volatility (i.e. the potential for incremental price changes to occur in the underlying asset in this case APPL, which may benefit the option holder prior to the options expiry). This we formally call the *time value* over the remaining 'life' of the option. In the case of such a call option on a stock (recollecting it represents *the right, but not the obligation to buy*), we could say that the Black-Scholes model is a pricing formula that equals the minimum value - being the Stock price minus the Exercise (*or Strike)* price; i.e. the 'intrinsic' value part plus the time value, i.e. the positive value of any anticipated volatility in the stock price (either implied from the pricing of that or similar options, or as actually calculated from the observed standard deviation of the stock as measured over a suitable historical time period, say a few years).

In the case of a dividend paying stock (that becomes ex-dividend or pays out during the option's life), however, the option owner does not receive any benefit as the stock is not owned until after the option is exercised, so we must subtract the value of any dividend paid, (as the then current owner of the underlying stock will *still be in possession and thus entitled to receive any such dividend)* until the option is *exercised*). This necessarily lowers the underlying value of the stock post *ex-dividend* date and hence the option value assuming it was not previously exercised.

Figure 1: Call option values versus stock price



In essence at the expiry of option, (in the case of a European Option, or at exercise in the case of an American Option, the payoff is simply the difference between the price at which you have the right (*but not the obligation*) to buy the underlying asset at, and the current market price of the asset. At this point it is usually appropriate to explain the difference between an American and European style option which depends on when it can be exercised, and to reinforce that they might consider that Americans are not so easily satisfied, and thus always demand more than Europeans; hence the price of an 'American' option is said to always be greater than, or to 'dominate', the value of a European option prior to expiry, given the inherent flexibility to be exercised ahead of time rather than solely at expiry. To recap therefore:

1. We have started with the minimum value (i.e. the intrinsic, or lower bound) of a call option what it is worth at any given moment: In words, the call value equals the stock price minus the discounted present value of the underlying strike price (this difference being known as the *intrinsic value*) or C = S – PV (K). The present value is taken of K because if we do not exercise immediately, the monies required to provide the strike price K on exercise may remain invested, potentially growing continuously at a riskless interest rate *r*. This is because we can invest the proceeds required upon exercise at some deposit rate *r*, in a bank for example. Hence, the present value of the future intrinsic value will be worth exactly this minimum value. Indeed, before the Financial Accounting Standards Boards accounting rule, known as FASB123R, this was the value that private US companies could use for valuing long-dated employee options in their accounting statements. This was because these employees have no traded volatility to contend with (i.e. there was no time value in the absence of trading) – since employees, are unable to 'trade' their corporate option entitlements prior to vesting).
2. Now all we need to do is wrap this minimum value in combination with two cumulative normal distribution functions: $N(d_1)$ - regarding all potential outcomes for the future stock price S in excess of the strike price, and N(d2) the probability of the stock price S being at least equal to the value of the known *a priori* exercise price K of the option. These distributions are simply used as a statistical measure or proxy for the range of possible stock price outcomes between now and the options expiry. These are based on a standard Normal (*or Gaussian*) distribution, the moments of this distribution can readily be estimated from historical asset pricing, in terms of their mean μ and standard deviation σ.
3. At the maturity time of the option T, the Call price thus becomes: $C= S*N(d_1) – PV(K)*N(d_2)$
   … which is of course the famous Black-Scholes formula.

At This Point Most Students Still Ask Why?

Because essentially in the case of a call option, it is worth the difference between the *uncertain* stock price in the future, and the *known* exercise price today, discounted by the uncertain time period that will elapse between now and whenever it is exercised (or expires). During this uncertain time interval, the prevailing interest rate will be applied to discount the value of the known future exercise price.

Our mnemonic of 'just add on the value of the volatility' to the intrinsic value of the option, is a simply way to teach students to memorize the time value and normal distribution elements embedded in the Black-Scholes formula. Thus we take the minimum intrinsic value of the option, based on the difference between today's price and the exercise price, and simply "wrap in" the N(d1) and N(d2) probabilities of positive price changes into it (negative price changes cause the option to expire worthless). Together these probabilities increase the minimum intrinsic value to a higher price which includes the time value element: as the more volatility (and time) of course, the more potential value will be added to the option's price. It should not be forgotten that the intrinsic value can never be negative, even if the asset price is below the strike or exercise

price. In such case, given the right but *not* the obligation to exercise, an option with a negative intrinsic value at expiry, is simply discarded as worthless.

In effect then this is just a memory technique for students, but it also has a worthy theoretical justification, for if the volatility were indeed zero, then you would arrive at the intrinsic value (as if volatility is zero then the price is completely stable and thus the future value known for certain) - which again makes the formula nothing more than a simple present value calculation, albeit using continuous as oppose to simple compounding, whose values is the difference between the strike and the future exercise price.

Digging Deeper - What are the N()s in Black-Scholes?

The N()s are simply the cumulative normal (distribution) functions (meaning the percentage probability of an occurrence, or simply the area bounded by the observed value and the lower asymptotic limit $-\infty$ on a plot of the standard normal curve). N(1) = 84%, because one standard deviation (i.e. $+1\sigma$) has 84% of the area under a standard normal curve to the left (remembering from statistics that although approximately 68% of the standard normal curve is captured within one standard deviation, it is necessary to add half again of the residual area, i.e. half of the remaining 32%, due to the area contained beyond one standard deviation shown within the negative asymptotic tail).

Together, the N(d1) and N(d2) simply probability-adjust the resulting pricing formula.

$N(d_2)$ is the probability that the call will be exercised; i.e. that it will finish in the money or that the Stock price will exceed the Exercise price; as otherwise the call option is worthless and will be discarded if the exercise price exceeds the stock price at expiry. In this sense $KN(d_2)$ is the strike price multiplied by the probability that this strike price will in fact have to be paid; and where $S*N(d_1)\exp(-rt)$ is the discounted expected value of the future stock price, if and only if the call option subsequently finishes in the money - the difference in such a case being the value of the option at exercise. You can see these effects in action using the following web link, for example: http://www.editgrid.com/bt/frm2007/LO_34.x_BlackScholes

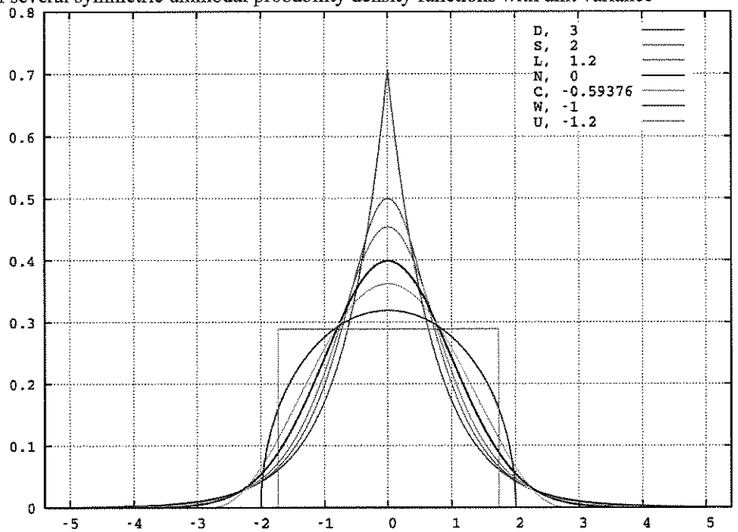
Meissner (2007) further demystifies $N(d_2)$, the probability of exercising a call and N(-d2) the probability of not exercising a call (or of exercising a put) as well as providing a good explanation, as to why $N(d1) \geq N(d2)$, which helps to see why "adding volatility" to a minimum value (or lower intrinsic bound) is always an addition in terms of value to a European style call option. But what is even more relevant in this article is the centrality and generality of the Black-Scholes model for both option pricing (of derivatives) and the Merton model that is used in assessing credit risk. 'The 1974 Merton corporate model [i.e. one that treats equity as a call option on the firm's underlying assets (and analogous to the stock price because shareholders, collectively, have the 'right to own' the entire company by paying off any outstanding debt due to bond holders) with firm's debt as analogous to its strike price] is mathematically identical to the famous Black-Scholes-Merton 1973 option model".

At this point it should be recalled that the volatility of the asset price has been expressed based on the simplifying assumptions of the normal distribution (the N(d)), being the *normality assumption* in respect of the distribution of future asset prices as based on the central limit theorem. Students typically recall that this implies, that given a large enough sample size (or as n tends to infinity) sample distributions are expected to converge towards the Normal.

For theoretical reasons, under this central limit theorem, any variable that is the sum of a large number of independent factors is *likely* to be normally distributed also. For this reason, the normal distribution is used throughout statistics, natural science, and social science as a simple model to help explain the behavior of complex phenomena. Stock prices, being affected by a myriad of factors could be assumed to conform to this as they are in the Black-Scholes model. It is however important to mention that there are some valid concerns regarding the normal distribution approximation. Whilst undoubtedly an elegant, mathematically tractable and useful statistical device for handling uncertainty, it may not be entirely reflective of the underlying distribution of observed asset returns *(bearing in mind that a finite sample of an infinite population still need not necessarily conform to 'normal').* Whereas a standard normal is described entirely by its first two moments, i.e. its mean and standard deviation, financial data may not be, as a plot of share price returns may not necessarily empirically conform to that of a standard normal distribution or bell shaped curve, at least at the finite sample level. Indeed higher moments are typically expressed in financial data series. In fact financial data is often skewed and peaked in nature, and more readily described by a log-normal due to its asymmetry due in part to limited liability and skewed due to the positive drift implied by inflation and economic growth. Limited liability implies no more than 100% of an equity investors capital may be lost, so some positive skewness, otherwise known as the third moment typically affects the underlying distribution, which may cause a deviation from the Normal; since the distribution is not centered on plus or minus infinity, but truncated from -100% and plus infinity, and is consequently said to 'lean' or express positive skewness. In fact even with normalized distributions the median, mode and mean will be shifted by a mean pushed by extreme positive values. The study of the relationship between these three parameters will define the level of skewness apparent in financial data and is an important area of research. This skewness will be coupled with a more peaked nature to the distribution as well. This 'peakedness' being statistically referred to as the fourth moment of the distribution, is known as kurtosis, which is a measure of how peaked or flattened a plot of the

asset price distribution is relative to the standard Normal distribution. Financial data tends to be more highly peaked relative to a normal distribution and this is known as being *lepto*-kurtotic. The opposite phenomena is known as *platy*-kurtotic which describes a distribution that is flatter in appearance that the Normal.

Figure 2: A plot of several symmetric unimodal probability density functions with unit variance



From highest to lowest peak with embedded links to allow interested students further research:

- The D line, kurtosis 3, Laplace (D)ouble exponential distribution
- S, kurtosis 2, hyperbolic (S)ecant distribution
- L, kurtosis 1.2, (L)ogistic distribution
- N, kurtosis 0, The (N)ormal distribution
- C, kurtosis −0.593762…, raised (C)osine distribution
- W, kurtosis −1, (W)igner semicircle distribution
- U, kurtosis −1.2, (U)niform distribution

In fact, *lepto*-kurtosis causes more of the distribution to be found in the 'peak' and the 'tails'. This could reflect a number of market factors such as market microstructure, herding and momentum effects as well as transaction costs and bid-ask spreads. Moreover, this has a disproportionate effect on the area under the curve contained with respect to each successive standard deviation being calculated from the mean, versus a normal distribution. This alludes to why 'Black Swan' author Nassim Nicholas Taleb suggests that outlier events appear rather more frequently than the Normal distribution might otherwise predict particularly in respect of financial data. Statisticians and theoretical physicists have also looked at developing *lepto-kurtotic* distributions directly, such as the more promising *hyperbolic secant* and *logistic* distributions, but their higher moments - whilst nearer to modeling actual empirical data - do not appear to justify the additional computational effort – operational and computing power may of course refine this view in the future.

By their nature businesses profit-maximize and therefore are not normally distributed. Indeed, this comparatively peaked nature in respect of the underlying distribution of financial market data appears ideally suited to wrong foot market participants, such as fund managers or traders, as it implies periods of more relative stability in terms of financial returns, as more of the distribution is concentrated in the peaks than would be expected under the standard normality assumptions. This is followed by discrete and more distant outliers derived from the tails of the distribution – perhaps occurring at intervals long enough apart to catch out the majority of successive generations of market participants, lulling them into a false sense of security through '*pseudo*-predictability' or relative stability in returns. This could allow greed (manifested through excessive leverage and 'insurance' or moral hazard) to magnify the resulting apparently 'anomalous' outlier events - often with catastrophic financial consequences as witnessed in financial markets during the recent crisis. In part this is well recognised in option pricing as a limitation on the Black Scholes model and may explain why volatility levels from implied option prices, such as the VIX fear gauge (the measure of volatility implied from option prices on the S&P 500 constituents) appears to jump around quite a lot with successive periods of switching between high and low volatility amid periods of apparent

calm or stationarity in volatility [A technique known as a Generalized Auto Regressive Conditional Heteroskedasticity (or GARCH) process used to model these volatility shifts – see Hamilton (1994)].

One empirical solution to the problem of 'fat' tail risk that is proffered by the author Nassim Taleb (2007), is to use repeated Monte Carlo simulation (which is the random generation of successive series of data, based on the reordering of observations taken from an underlying historic data series, either with re-sampling of the observations or without depending on time scale as there may not enough observations without sampling. However this still requires *a priori* decisions on the appropriate length of underlying time series that should be used, with the commensurate validation questions of adding 'noise' versus 'information' in extending any financial time series. In the case of stock return data this requires careful consideration, given that firms themselves change in nature over time. An example could be the century old blue chip IBM, which as a company operating in 2012 is not necessarily reflective of the same set of risk factors that might have affected the original IBM business in 1912 - even though by nomenclature it is the same company.

The simplifying assumptions underlying the Black-Scholes model include: no transaction costs (or the bid-ask spreads evident in buy or selling either the asset or the option on the asset – clearly this assumption may be violated in the real world by what is known as liquidity risk, even in the most intensely competitive marketplaces. That the distribution of asset returns follows a log normal distribution – this is violated if there appears to be more extreme moves in practice than the normality assumption ought to predict – this is known as tail risk as associated with the difference between observations in the tail of the underlying distribution and that predicted by a normal curve. That the log normal process generating these asset returns is strictly stationary in terms of constant volatility – if however volatility tends to jump around, market participants will be exposed to volatility risk. That markets are open continuously for trading – the effect of opening and closing of markets around the world will expose participants to gap risk on the opening and closings prices across difference markets due to discrete opening hours. That there is one borrowing and lending interest rate that is known, constant and continuous – clearly, interest rates vary over time and borrowers and lenders will face differing interest rates due to bank intermediation margins.

Black-Scholes Revisited Intuitively

The easiest way therefore to understand the Black-Scholes formula intuitively is to consider what happens if an investor were to exercise the option. This has two parts:

i) Value of the cash to Buy the Option. Firstly, if the option is exercised, the owner must pay the strike price; but this is true only if the underlying stock price is above the strike at maturity. So to work out the expected value of this we simply need the probability that the stock price is above the strike price at maturity. Let's call this probability $N(d_2)$, and the strike price, K. Then the expected value of this is just $KN(d_2)$ which is the value of the cash flow at maturity. As discussed above, to get the value of this cash flow today we would need to discount it, and we would do so using the continuous discount factor is e^{-rT}. The exponent function simply results from the idea of continuously compounded returns r (negative in sign because we are receiving, not paying the return), over time T. Discrete compounding is observed on bank accounts, paying daily, weekly, monthly or annual interest – continual compounding merely extends this analogy by breaking a day into an infinite number of sub-periods, e.g. hours into minutes, seconds or even infinitely miniscule parts thereof, and assuming that interest can still be paid in commensurately infinitesimally but divisible amounts. So the value of the cash to buy the option today PV (K*N(d2) becomes $KN(d_2)\ e^{-rT}$ or the strike price, discounted continuously, and multiplied by the probability of exercise (or that the stock is at least as high as the strike price - since otherwise the option is not exercised and expires worthless).

ii) Value of the Stock Received, if any. Secondly, if the option is exercised we then get a unit of the stock. This is clearly worth whatever the stock price is in the market at maturity. Nevertheless, this again only happens *if the underlying stock price is above the strike at maturity* (as unless *obliged* to do so, investors do not rationally or willingly give up any money, altruistically). It turns out that the expected value of this, valued as at today, is in fact proportional to S, the stock price today, and can be written as $SN(d_1)$. That is to say, $SN(d_1)$ is the expected value of something that is equal to the final stock price, *if the final stock price is above the strike*, and equal to zero *if the final stock price is below the strike.*

Volatility (sigma, or σ)

If you have followed our discussion about options you will probably already be aware that their values depend on something called volatility. Volatility is usually not needed to price derivatives that are not options. Technically volatility is defined as the annualized standard deviation of the return on an asset (in our case an example, Apple stock) over a predetermined period e.g. 1 year. They are simply expressed as percentages. However, it is easier to think of it intuitively as the amount that the price will be expected to swing around an average price over any given period of time. Stocks with a high level of uncertainty surrounding them will therefore necessarily have higher volatilities. An example currently might be the

stock of small mining companies. Stocks that are relatively stable (e.g. AT&T or Wal-Mart whose businesses are highly stable) will necessarily have lower volatilities. In part, this volatility reflects the inherent uncertainty of earnings that such companies are exposed to in the market place.

Why Does Volatility Affect the Price of an Option?

Again, this is because our payoff graph is not symmetrical. A stock that has a high volatility is more likely to fluctuate or swing around more, and hence is more likely to have a very high value (or indeed a commensurately very low value) at maturity. A stock with a low volatility is more likely to be close to its current value at maturity. Now if the stock price at maturity is below our strike price, we do not care if it is just slightly below or indeed massively below. In both cases we simply do not exercise the option and it expires worthless (as we are not obliged to lose money) and so we do not make any money. But conversely if the share price at maturity is above our strike, ideally *we really want it to be as far above the strike price as is possible*, since we make more money the higher the corresponding value is at maturity.

Therefore, an option with a high volatility is more likely to make us lots of money if the price goes up, but will not lose us lots of money even if the price goes down hugely – as we can abandon the option for free. Hence there is a clear beneficial asymmetry explaining why *ceteris paribus…* options on underlying securities with high volatility are more valuable (i.e. desirable) than options on underlying with low volatility. As we will see below, both d1 and d2 in the values discussed above depend on volatility.

Black-Scholes Formula Revisited in Terms of d_1 & d_2

If an investor has bought the call option he is paying the cash amount in i) above and receiving the value of the stock ii). So we can say that the value, C, of a European call option on a non-dividend paying stock assuming continuous compounding is: $C= SN(d_1) -Ke^{-rT}N(d_2)$. So what are d_1 & d_2? As mentioned in the introduction the mathematics behind the calculation of the probabilities in the Black-Scholes formula are fairly complex. It turns out that if N() is the cumulative normal density function (a statistical operator based on the normal distribution) then d_1 and d_2 can be expressed as below. Note that whilst these formulas appear complicated, students and traders can just plug in the underlying values and get a result known as a 'closed form' solution. Where

$$d_1 = \frac{\ln\left(\frac{S}{K}\right)+\left(r+\frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}} \text{ and } d_2 = \frac{\ln\left(\frac{S}{K}\right)+\left(r-\frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}$$

where sigma *(σ)* is volatility, *(r)* is interest rate, *S* is the Stock Price, *K* is the strike price, and *T* is time.

It can readily be observed that d_1 must always be greater than d_2, providing $T > 0$ i.e. there is some time left prior to expiry given that $d_2 = d_1 - \sigma\sqrt{T}$. The 'Greeks' or the sensitivity of the value of an option is based on changes in underlying input parameters. The first Greek is called Delta and simply measures the sensitivity of the price of the option to changes in the underlying price of the Stock (or other asset) from which the option's value is being derived. In practice, this is the most important parameter monitored in any trading environment since it tells a trader (and more importantly the risk manager) how much price risk exposure is being taken in trading the option position.

N(d_1) versus N(d_2); or Option Delta Versus the Probability To Exercise

The delta of an option, its responsiveness to changes in the underlying asset price, is frequently considered to be the same as the probability that an option will be exercised, i.e., the probability that the option will be in the money at maturity. However there is a difference which is especially important when it comes to long-dated options on volatile stocks. Consider a Simple Binomial Tree. Assume the stock price today is $100 and it will be either $150 or $50 when the European call option expires (a one-step binominal tree) with 50% probability respectively, that is:

If strike price=$100.00, then the probability to exercise =0.5 & Delta= $\frac{((150-100)-0)}{(150-50)} = 0.5$.

If strike price=$120.00, then the probability to exercise =0.5 & Delta= $\frac{((150-120)-0)}{(150-50)} = 0.3$.

If strike price = $149.99, then the probability to exercise=0.5 & Delta= $\frac{((150-149.99)-0)}{(150-50)} = 0$ (approximately).

This simple example shows that delta and the probability to exercises are in fact different. More specifically, option delta - the hedging ratio - does not only care about the probability of the option ending up in the money, but it also cares about how deep the option is in, or out, of the money - as the final payoff of the option depends on where spot is in relation to the strike

price. N(d1) is not simply a probability, it is in fact delta (for a European call on non-dividend paying stock). And N(d2) is the probability that the European option *expires* in the money.

So What is Delta (δ)?

Option delta and the probability to exercise are also distinguished in the Black-Scholes formula. Recall that the pricing formula for a European call is: Call option price c = S0*N(d1) – PV(K)*N(d2), where: S_0 = Current Stock Price, PV(K) = Present Value of Strike Price K, N(d1) and N(d2) are the cumulative probability distribution functions for a normal distribution (i.e. it is the probability that such a normal variable will be less than d_1 or d_2 respectively) and $d_2= d_1- \sigma\sqrt{T}$ (σ is of course the stock volatility as measured by its standard deviation and T is the time to maturity).

By definition, we immediately have N(d_1) as the option delta, representing the changing rate of the option price as a result of the stock price change. It can be further shown that N(d_2) actually is the probability the option will be exercised. Since d_1 is always larger than d_2 (as long as there is some time remaining to expiry), it follows the cumulative probability function N(d_1) - the option delta - should always be larger than N(d_2) - the probability to exercise. Further, because the difference between d_1 and d_2 is $\sigma\sqrt{T}$ (or σ*T^1/2) such a difference will be more significant for long dated options (with a large T) on highly volatile equity stocks (with a larger σ), hence the difference between the option delta and probability to exercise, and why a longer time to expiry and a higher volatility always raise option prices.

An Intuitive Explanation

Instead of digging into mathematical calculations, we are able to derive an intuitive explanation from the Black-Scholes formula, if we just accept that N(d_1) is the option delta and N(d_2) is the probability to exercise. Suppose for a moment, N(d_1) is the same as N(d_2), the exercise probability. Then, the value of a European call option would be: Value = (Probability of Exercise) * (S - PV(K)). Now suppose the option is out-of-the-money, but there is still time to maturity. The present value, S - PV(K), of the expected exercise pay-off could well be negative, which would result in a negative option value. This cannot be correct, since an option can never have a negative value (due to downside insurance). This explains why N(d_1) needs to be larger than N(d_2); so the option value will not be negative when the stock price drop below the strike price.

The reason for this is that the above formula acts as if it was known what the option holder will get when the option is exercised, and the only remaining uncertainty is if the option will in fact be exercised. Nevertheless, this is clearly not the case: there is uncertainty about the stock price at maturity and the higher the volatility, the longer the option life, the more stock price uncertainty there is. Due to the limited liability/downside implicit in an option, higher uncertainty can only be beneficial (since we never exercise options with non-positive intrinsic value, as they are freely discarded at no cost at all by the holder), so that the value formula of the option needs to be adjusted to reflect this extra value, which is reflected in the excess of N(d_1) over N(d_2). Technically, option pricing theory will attribute the difference between N(d_1), the option delta and N(d_2), the probability to exercise, to probability calculations in different "measures" - mathematics-oriented student being referred to Bjork (2008) or similar publications.

Some Delta Conclusions

The delta of an option is not necessarily equal to the probability to exercise the option. In general, for T>0 the option delta is larger than the probability of exercise, and the difference becomes more significant with respect to long dated options on more volatile equity stocks. That is because delta incorporates not only the probability the option will be exercised, but also the amount by which the option is in the money. Even an option that is heavily in the money already still has a probability of rising further whilst there is some residual time to expiry. The Black-Scholes formula also shows that delta has to be adjusted by more than the probability to exercise in order to reflect the option's value increase because of stock price uncertainties, and such adjustment becomes more significant for long dated options on volatile equity stocks.

Some Volatility Extensions – Vega (ν)

The impact of volatility is known as Vega (i.e. the change in the option price, for any given change in underlying volatility). Vega approaches zero as the option goes, in both directions, in and out-of-the money. Intuitively, if the option is far out of the money, adding volatility just doesn't help you in any reasonable timeframe (e.g. you would need say 1 year plus – but traded options run out in at most nine months). Conversely if your option is deep in the money, you are already near the value of the underlying asset (i.e. the share price itself) and the option can't get to be worth any more than its underlying (as

the Stock price becomes significantly greater than the strike price, or indeed any multiple thereof, you are already effectively very close to the upper value bound which is S.

Collectively, the $N(d_1)$ and $N(d_2)$ are probability-adjusting the minimum intrinsic value, where minimum value (MV) = Stock - (Strike)EXP[(-rate)(t)]. That MV = Black-Scholes value in the special case where the volatility is zero [if volatility = 0, then of course both $N(d_1)$ and $N(d_2)$ = 1]. Now, as we "add volatility" the spread between $N(d_1)$ and $N(d_2)$ increases (i.e., $N(d_1)$ - $N(d_2)$ which we might think of as a volatility wrapper around the MV) such that the Black-Scholes value increases with greater volatility. Hull 2009 (see Chapter 13 in particular) provides another interpretation, based on rearranging e.g., $N(d_2)$ is the probability that the option will be struck, i.e. that it will at least reach the exercise price.

Other Greek Sensitivities:

Rho (ρ) or the sensitivity of the option price to changes in the interest rate (r). This affects the present value elements of the formula, making call options inherently less valuable as interest rates rise; vice versa for puts. Theta (θ) is the Greek measuring the sensitivity of the option price to time. This is typically based on a calendar (365) or trading day (252) as a fraction of a year. Increasing time always increases option prices, as this allows for far greater volatility of returns in the underlying asset. Gamma (γ) is the sensitivity of delta (δ) to changes in the underlying asset price, or the second derivative of price. In similar fashion, this alludes to convexity the second derivative versus duration the first derivative, as used in measuring the sensitivity of bond prices to interest rates. It should be recognised that some sensitivities are usually quoted in scaled-down terms, to match the scale of likely changes in the parameters. For example, rho is often reported divided by 10,000 (1 basis point rate change, where 1 basis point is 100th of 1%), Vega by 100 (1 volatility point or 1% change), and theta by 365 or 252 (1 day decay) based on either calendar days or trading days per year).

Conclusion - No Actual Derivation of the Black-Scholes Formula

This note has attempted to provide an intuitive interpretation of the Black-Scholes formula volatility, without going into the mathematics behind it. Such an interpretation inevitably glosses over some of the technical details such as risk neutrality considerations, for instance (i.e. investors are indifferent to a $1 for sure versus and gamble that has an expected value of $1). In particular, the actual derivation of the Black-Scholes formula, of course, was not done directly using the intuitive ideas that have been just discussed here and the interested reader is referred to Merton (1992) for a rigorous mathematical proof.

References

Bjork, Tomas. 2008. *Arbitrage Theory in Continuous Time*, Oxford University Press, USA, 3rd Edition.

Black, Fischer, Myron Scholes. 1973. 'The Pricing of Options and Corporate Liabilities', *Journal of Political Economy,* Vol. 81, Issue 3, pp.637-654.

Chen, Jing. 2010. 'An Intuitive Understanding of the Black-Scholes Formulas' Working Paper, September 24th. Available at SSRN: http://ssrn.com/abstract=1682261 or doi:10.2139/ssrn.1682261

Datar, Vinay T., Scott H. Mathews, 'European Real Options: An Intuitive Algorithm for the Black-Scholes Formula', *Journal of Applied Finance*, Vol. 14, No. 1, Spring/Summer 2004. Available at SSRN: http://ssrn.com/abstract=560982

Hamilton, James D. 1994. *Time Series Analysis*, Princeton University Press, 1st Edition.

Hull, John C. 2009. *Options Futures & Other Derivatives*, Prentice Hall, 8th Edition.

Meissner, Gunter. 2007. Probability of Default, *GARP Risk Review*, Global Association of Risk Professionals, Issue 4, July/August.

Merton, Robert C. 1973. 'Theory of Rational Option Pricing', *Bell Journal of Economics and Management Science,* Volume 4 (1), pp141–183.

Merton, Robert C. 1974. "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates," *Journal of Finance*, American Finance Association, vol. 29(2), pages 449-70, May.

Merton, Robert C. 1992. *Continuous-Time Finance*, Blackwell Publications, Revised Edition.

Taleb, Nassim N. 2007. *The Black Swan: The Impact of the Highly Improbable*, Random House.

The U.S. Stock Market Reactions to the WSJ Daily Stock Picking

Sung C. No and Michael Smyser, Southern University and A&M College
Doh-Khul Kim, North Central College

Abstract

This paper examines the Wall Street Journal's daily choice of four "Good stocks" and four "Bad Stocks" selected by a group of the WSJ stock analysts. The current study initially identifies 2,828 stocks in the data spanning from December 3, 2007 to December 31, 2009. Using the Brown-Warner event study methodology, the study found statistically significant price responses on one day after the publication day. Moreover, the study suggests that most of these significant returns be generated by repeated stocks or most noteworthy stocks, which is consistent with the investor overreaction hypothesis. The study also found that the responses of good and bad stocks are asymmetric across exchanges.

Introduction

In the Tuesday thru Friday editions of *The Wall Street Journal* (WSJ) is published a column titled "Money & Investing" that highlights a selection of eight common stocks: four stocks associated with "Good News" and four stocks associated with "Bad News". The stocks chosen for the article highlight the most newsworthy companies based on the previous day's price action. WSJ analysts employ multiple criteria to choose the four stocks in the "Good News" (hereafter, Good Stocks) and four stocks in the "Bad News" (hereafter, Bad Stocks). First, the choice set usually includes any large-cap, extremely newsworthy stocks that are the "story of day" in the paper's stock market summary. However, if the price change of a candidate large-cap stock is not at least 2-percent, the stock is not typically included in the column. Second, the WSJ analysts consider stocks that have moved 5-percent or more and select ones that can provide the most interesting story to WSJ readers. Third, the Journal editors tend to get a good mix of financial and tech stocks.

For example, if three money-management stocks had large price changes due to an analysts' report, the editor typically includes only the biggest mover, with the other two stocks mentioned in commentary placed above the price chart displayed for each of the eight chosen stocks. Finally, the journal's stock pickers tend to avoid microcap stocks that had a large percentage price change but relatively small in absolute dollar terms (e-mail communication with a WSJ analysis, 2010).

Numerous previous event studies have examined stock price reactions to the release of price information (Yau, *et al.*, 1994: stock price reactions to events reported in the WSJ; Ferreira and Smith, 1999: stock price reactions to small stock focus column; Mathur and Waheed, 1995: stock price reactions to Business Week's Inside Wall Street column; Conrad, Cornell and Landsman, 2002: stock price responses to good and bad earnings shock among others). However, no research has yet to examine whether abnormal returns can be earned by conditioning trades on the eight stocks selected by the WSJ in the "Money and Investing" column. The purpose of this paper is to determine if the portfolio of four Good Stocks can beat the market by consistently yielding positive abnormal returns and whether the portfolio of four Bad Stocks consistently produces negative abnormal returns.

Data and Descriptive Statistics

The study compiled the eight stocks presented at the daily column of the "Stocks in the News" that appears in Section C of the *Wall Street Journal*. The data sample begins December 3, 2007 and end December 31, 2009. An initial sample of 2,828 stocks was identified by searching the daily issues of the WSJ. Observations were retained only if the firm associated with the announcement had sufficient returns on the Center for Research Security Prices (CRSP) daily tapes.

The number of identified observations by CRSP is 1,868 and 1,853 for stocks in Good News and in Bad News, respectively. The WSJ analysts select stocks from various exchanges, such as the NYSE/AMEX, NASDAQ, OTC, LSE, Frankfurt, TSE, Toronto, TLV, among others. Each observation in this paper is classified into one of three exchange groups: the NYSE, NASDAQ, or OTH (for other) which resulted in a composition of 62.69%, 29.01%, and 7.7%, respectively for the "Good Stocks" sample and a composition of 64%, 27.31%, and 8.58%, respectively for the "Bad Stocks" sample.

On the day prior to appearance in "Stocks in the News" column, the average rate of return for "bad stocks" in the news was -12.79% while the average return for "good stocks" in the news was 15.03%. Including many repeated stocks, a total of 1,022 stocks were selected for good stocks with General Motors Co. appearing most frequently (34 times). In contrast, Citigroup, Inc. was selected 33 times as a stock in the Bad News during the sample period. Out of 2,224 observations, only

982 different stocks were chosen as stocks in the bad news due to repeated stocks. By exchange, 1,386 of good stocks are listed at the NYSE; 667 at the NASDAQ; and 171 of them at the other exchanges, such as TSE, LSE, Frankfurt stock exchange among others. On the other hand, 1,415 of bad stocks are listed at the NYSE; 620 at the NASDAQ; and 189 at the other exchanges.

Methodology

To examine whether or not four good stocks outperform the average stock in the market after the publication date, abnormal returns are calculated using the standard market model event study approach in Brown and Warner (1985). Daily expected residual returns are calculated for each transaction and averaged across firms. An 11 day window is used, centered on the event-day zero defined as the publication date that the eight stocks appear. Prediction errors were determined based on the market model.

Specifically, for each stock *i*, abnormal returns are calculated on each day during the interval from t = -5 through t= +5 using the two step procedures below: First, the market model is estimated by ordinary least squares:

$$R_{i,t} = \beta_0 + \beta_1 R_{m,t} + \varepsilon_{i,t} \tag{1}$$

where $R_{i,t}$ and $R_{m,t}$ are, respectively, the rate of return on stock i and the rate of return on the CRSP equally weighted index on day t; $\varepsilon_{i,t}$ is the disturbance term; and β_0 and β_1 are parameters to be estimated. The market model (1) is estimated using returns from a sample beginning at t = -244 and ending at t = -6 relative to the event day t=0. Abnormal returns are then calculated for each stock on each day during the 11-day window from time t = -5 through time t = +5 as follows:

$$AR_{i,t} = R_{i,t} - \hat{\beta}_0 - \hat{\beta}_1 R_{m,t} \tag{2}$$

where $R_{i,t}$ is the rate of return on i^{th} stock on day t, and β_0 and β_1 are parameter estimates for stock i from (1). The estimation period used here permits the identification of abnormal price behavior relative to the one year period prior to the 11-day window centered on the event-day.

In the absence of an abnormal performance, the expectation of the abnormal return (2) is zero. Testing whether the return on day t is abnormal is facilitated by constructing the standardized abnormal return (SAR):

$$\overline{SAR_t} = \overline{AR_t} / \hat{S}(\overline{AR_t}) \tag{3}$$

where

$$\overline{AR_t} = \frac{1}{N_t}\sum_{i=1}^{N_t} AR_{i,t} \tag{4}$$

$$\hat{S}(\overline{AR_t}) = \sqrt{\left(\sum_{t=-244}^{t=-6} (\overline{AR_t} - \overline{\overline{AR}})^2\right) / 238} \tag{5}$$

$$\overline{\overline{AR}} = \frac{1}{239}\sum_{t=-244}^{t=-6} \overline{AR_t} \tag{6}$$

A test statistics Z_t is constructed as:

$$Z_t = (N)^{-0.5} * \sum_{i=1}^{N_t} SAR_{i,t} \tag{7}$$

where N_t is the number of sample stocks whose abnormal returns are available at day t. If the standardized abnormal returns (the $SAR_{i,t}$'s) are independent and identically distributed with finite variance, in the absence of abnormal performance, the test statistic will be distributed unit normal for large N_t. An average cumulative abnormal return for days T_1 through T_5 is calculated as:

$$\overline{CAR}_{T_1,T_5} = \sum_{t=T_1}^{T_5} \overline{AR_t} \tag{8}$$

In the absence of an abnormal performance, the expectation of the cumulative abnormal return (8) is zero. Testing whether the cumulative returns from day T_1 = +1 through T_5 = +5 are abnormal is facilitated by constructing the standard deviation of the cumulative abnormal return on stock i from day T_1 through day T_5. For tests over the (+1, +5) interval, the test statistic is the ratio of the cumulative mean abnormal return to its estimated standard deviation, and is given by

$$\overline{SCAR} = \sum_{t=+1}^{+5} \overline{AR}_t \Big/ \left(\sum_{t=+1}^{+5} \hat{S}^2(\overline{AR}_t)\right)^{\frac{1}{2}} \tag{9}$$

The test statistic $Z(T_1,T_5)$ is constructed as:

$$Z(T_1,T_5) = N^{-0.5} * \sum_{i=1}^{N} [\sum_{t=T_1}^{T_5} AR_{i,t} / CAR_{i,T_1,T_5}] \tag{10}$$

If the assumptions presented in the discussion below of the day t abnormal returns are satisfied and if the abnormal returns are not serially correlated, then the test statistic $Z(T_1,T_5)$ is distributed standard normal under the null hypothesis. (Brown and Warner, 1985).

Empirical Results

Abnormal returns from the full sample

Table 1 displays the average abnormal return ($\overline{AR_t}$), the test statistic Z_t, the average cumulative abnormal return ($\overline{CAR_t}$), its test statistic, and the number of positive and negative observations for the full sample on each day during the 11-day interval centered on the event day (i.e. from t = - 5 through t = +5), Overall, the results indicate that stock prices are significantly impacted when those stocks are included with the WSJ publication of Good and Bad stocks. In particular, For the Good stocks, the average abnormal returns on the day preceding the publication day (t = -1), the publication or event day (t = 0), and the day following publication (t = +1) are, respectively, 13.91, 0.60, and 0.25% (test statistics Z_{-1} = 120, Z_0 =5.12, and Z_{+1} = 2.14). Note that $\overline{AR_t}$ is significantly positive on each of the three days prior to publication day (t = 0).

Table 1: Abnormal Returns

	Good Stocks				*Bad Stocks*			
Day	**N**	**AR**		**+ : -**	**N**	**AR**		**+ : -**
AR_{t-5}	1810	-0.04%	(-0.32)	864:946	1821	-0.26%	(-2.32)^b	855:966
AR_{t-4}	1810	0.11%	(0.93)	873:937	1821	-0.38%	(-3.43)^d	859:962
AR_{t-3}	1810	0.40%	(3.44)^d	906:904	1821	-0.57%	(-5.13)^c	843:978
AR_{t-2}	1809	0.38%	(3.25)^c	930:879	1819	-0.32%	(-2.88)^c	855:964
AR_{t-1}	1810	13.91%	(120)^d	1753:571	1821	-11.61%	(-104)^d	89:1732
$AR_{t=0}$	1810	0.60%	(5.12)^d	903:907	1821	-0.24%	(-2.16)^b	823:998
AR_{t+1}	1805	0.25%	(2.14)^b	877:928	1813	-0.19%	(-1.67)^a	827:986
AR_{t+2}	1802	-0.02%	(-0.15)	857:945	1808	-0.08%	(-0.72)	857:951
AR_{t+3}	1798	-0.25%	(-2.14)^b	846:952	1802	-0.15%	(-1.36)	839:963
AR_{t+4}	1798	-0.25%	(-2.14)^b	813:985	1801	-0.09%	(-0.80)	857:944
AR_{t+5}	1793	-0.24%	(-2.04)^b	804:989	1796	0.06%	(0.55)	820:976
$CAR_{t=-5,-2}$	1810	0.85%	(3.65)^d	948:862	1821	-1.53%	(-6.88)^c	831:990
$CAR_{t=-1,0}$	1810	14.51%	(88)^d	1706:104	1821	-11.85%	(-75)^c	161:1660
$CAR_{t=+1,+5}$	1806	-0.50%	(-1.92)^a	844:962	1813	-0.44%	(-1.78)^a	826:987

Note that the numbers in parenthesis are Z-value. The letters a, b, c, and d denote statistical significance at the 0.10, 0.05, 0.01 and 0.001 levels, respectively, using a two-tail test.

Of course, a large absolute return on the day prior to publication (t = -1) provides the major signal to the WSJ analysts that the stock is experiencing a significant event and, hence, the stock becomes a candidate for possible inclusion in the WSJ column. The significance of abnormal returns during multiple days prior to the publication day (t = -3 and t = -2) suggests that some investors may be sufficiently well informed to make purchases of Good stocks in advance of publication. The significance of the abnormal return on the day after publication (t = +1) is interpreted as investor overreaction in previous

event studies (Palmon, Sun, and Tang, 1994). The average cumulative abnormal return during the period t = -1 to 0 is 14.51% with test statistic Z (-1, 0) = 88. A form of market correction appears to occur after the initial overreaction at day t = +1 since the abnormal return is significantly negative on days t = +3, +4, +5. The relative number of positive versus negative returns in the sample suggests that the results should not be attributed to the presence of a few outliers.

Table 1 also shows results for the full sample of Bad stocks. The average abnormal returns one day before the publication day, the publication day, and the day following day for the Bad stocks are, respectively, -11.61, -0.24, and -0.19% (test statistics Z_{-1} = -104, Z_0 =-2.16,Z_{+1} =-1.67). All these excess returns are statistically significant. Similar to the results with Good stocks, the significance of abnormal returns prior to the publication day (t = -5, -4, -3, -2) suggests that some investors may be sufficiently well informed to make sales of Bad stocks in advance of publication. Indeed, the evidence in support of informed trading prior to publication is stronger with the Bad stock sample. The significance of the abnormal return on the day after publication (t = +1) can be interpreted as investor overreaction similar to the Good stocks. The average cumulative abnormal return during the period t = -1 to 0 is -11.85% [test statistic Z (-1, 0) =-75]. Overall, the abnormal returns from the sample of Good stocks are slightly higher than the absolute value of abnormal returns from the sample of Bad stocks.

Repeated inclusion of stocks as Good or Bad

Over the sample period, many stocks were selected more than once as a good stock or a bad stock or both because the stocks met the general selection criteria adopted by the WSJ analysts on multiple occasions. For example, GM stocks were labeled as a good stock 34 times from December 3, 2007 through June 8, 2009 when the company filed for bankruptcy. In contrast, due to chronic bad news, Citigroup was included 33 times with the Bad stocks.

To examine the affect of stocks repeating, samples were constructed by eliminating all multiple appearances by including in the sample only the first appearance of a particular stock. The results are reported in Table 2 for Good and Bad Stocks. About 1,000 repeated observations were trimmed in Table 2. Average abnormal returns one day prior to the publication, the publication day, and the day following publication are, respectively, 17.00, 0.37, and 0.46% (test statistics Z_{-1} = 109, Z_0 = 2.35,Z_{+1} = 2.96). The significance of these returns is stronger than those in the full sample. The cumulative average abnormal returns are also higher after deleting repetitions. Similar results are found for Bad Stocks in Table 2. The average abnormal return on the publication day and the day following publication are, respectively, -14.30 and -0.34 (test statistics Z_{-1} = -101, Z_0 =-2.37). The significance is again stronger than that reported in the full sample. The average cumulative abnormal return during the period t = -1 to 0 is -14.63% [test statistic Z (-1, 0) =-73].

Table 2: Abnormal Returns: Without Repeating

		Good Stocks				*Bad Stocks*		
Day	***N***	***AR***		***+ : -***	***N***	***AR***		***+ : -***
AR_{t-5}	855	-0.28%	(-1.79)[a]	431:424	838	-0.19%	(-1.34)	400:438
AR_{t-4}	855	-0.14%	(-0.88)	399:456	838	0.30%	(2.09)[b]	419:419
AR_{t-3}	855	0.11%	(0.68)	414:441	838	-0.37%	(-2.60)[c]	382:456
AR_{t-2}	855	-0.27%	(-1.71)[a]	419:436	837	-0.01%	(-0.05)	398:439
AR_{t-1}	855	17.00%	(109)[d]	844:11	838	-14.30%	(-101)[d]	18:820
$AR_{t=0}$	855	0.37%	(2.35)[b]	411:444	838	-0.34%	(-2.37)[b]	390:448
AR_{t+1}	852	0.46%	(2.96)[c]	422:430	833	0.03%	(0.21)	367:466
AR_{t+2}	851	0.06%	(0.38)	400:451	830	-0.16%	(-1.16)	383:447
AR_{t+3}	849	-0.10%	(-0.63)	412:437	828	-0.26%	(-1.81)[a]	388:440
AR_{t+4}	849	-0.11%	(-0.68)	376:473	828	-0.13%	(-0.91)	392:436
AR_{t+5}	849	-0.21%	(-1.32)	379:470	826	-0.10%	(-0.72)	378:448
$CAR_{t=-5,-2}$	855	-0.58%	(-1.85)[a]	429:426	838	-0.27%	(-0.95)	406:432
$CAR_{t=-1,0}$	855	17.37%	(79)[d]	824:31	838	-14.63%	(-73)d	46:792
$CAR_{t=+1,+5}$	853	0.11%	(0.32)	403:450	833	-0.62%	(-1.95)[a]	383:450

Note that the numbers in parenthesis are Z-value. The letters a, b, c, and d denote statistical significance at the 0.10, 0.05, 0.01 and 0.001 levels, respectively, using a two-tail test.

Samples were also constructed by dropping all observations if the stock appeared as a Good or Bad stock more than once. As reported in Table 3, for Good stocks that never repeat, the average abnormal return one day prior to the publication is far greater than returns in the full sample. However, on the publication day and one day after publication the abnormal returns are substantially lower than the full sample. Indeed, the trivial positive returns after the event day are not statistically different from zero. The study found a similar pattern with the abnormal negative returns for the sample of non-repeating Bad Stocks

in Table 3. Importantly, these results indicate that abnormal returns generated by repeated stocks or what can be considered as the "most note worthy" companies in the news are the primary cause of significance and which, in turn, is evidence in support of the investor overreaction hypothesis in the empirical financial literature.

Table 3: Abnormal Returns: Dropping Any Repeated Stocks

		Good Stocks				*Bad Stocks*		
Day	***N***	***AR***		***+ : -***	***N***	***AR***		***+ : -***
AR_{t-5}	461	-0.32%	(-1.38)	233:228	449	-0.19%	(-0.81)	216:233
AR_{t-4}	461	-0.16%	(-0.67)	216:245	449	0.28%	(1.19)	218:231
AR_{t-3}	461	0.20%	(0.86)	228:233	449	-0.16%	(-0.68)	199:250
AR_{t-2}	461	-0.20%	(-0.86)	227:234	448	-0.03%	(-0.15)	212:236
AR_{t-1}	461	20.65%	(89)[d]	456:5	449	-15.81%	(-67)[d]	5:444
$AR_{t=0}$	461	0.11%	(0.48)	208:253	449	-0.44%	(-1.86)[a]	213:236
AR_{t+1}	458	0.15%	(0.64)	230:228	445	-0.12%	(-0.51)	195:250
AR_{t+2}	457	0.02%	(0.07)	214:243	442	-0.15%	(-0.62)	204:238
AR_{t+3}	455	-0.16%	(-0.69)	213:242	440	-0.22%	(-0.94)	205:235
AR_{t+4}	455	-0.34%	(-1.48)	204:251	440	-0.11%	(-0.48)	200:240
AR_{t+5}	455	-0.13%	(-0.55)	231:242	438	-0.19%	(-0.80)	190:248
$CAR_{t=-5,-2}$	461	-0.48%	(-1.03)	231:230	449	-0.10%	(-0.22)	212:237
$CAR_{t=-1,0}$	461	20.76%	(63)[d]	444:17	449	-16.25%	(-48)[d]	18:431
$CAR_{t=+1,+5}$	459	-0.46%	(-0.89)	217:242	445	-0.79%	(-1.48)	209:236

Note that the numbers in parenthesis are Z-value. The letters a, b, c, and d denote statistical significance at the 0.10, 0.05, 0.01 and 0.001 levels, respectively, using a two-tail test.

Observations grouped by exchange

Several previous studies report that stock responses to event news might differ across exchanges. Thus, abnormal returns were recalculated for samples constructed by grouping stocks by exchange. Table 4 summarizes the findings for Good stocks traded on the NYSE. Average abnormal returns one day before publication, the publication day, and the day after publication are, respectively, 11.54, 0.59, and 0.25% (test statistics Z_{-1} = 77, Z_0 =3.96,Z_{+1} = 1.70). All the excess returns are statistically significant. The average cumulative abnormal return during the period t = -1 to 0 is 12.13% [test statistic Z (-1, 0) = 57].

Table 4: Abnormal Returns for Good Stocks across Exchanges

		NYSE				*OTHERS*		
Day	***N***	***AR***		***+ : -***	***N***	***AR***		***+ : -***
AR_{t-5}	1135	-0.16%	(-1.06)	533:602	141	-1.07%	(2.54)[b]	62:79
AR_{t-4}	1135	0.12%	(0.80)	562:573	141	-1.75%	(-4.14)[c]	59:82
AR_{t-3}	1135	0.70%	(4.65)[d]	577:558	141	-1.05%	(2.48)[b]	65:76
AR_{t-2}	1135	0.18%	(1.19)	580:555	141	-0.16%	(-0.39)	72:69
AR_{t-1}	1135	11.54%	(77)[d]	1092:43	141	19.83%	(47)[d]	138:3
$AR_{t=0}$	1135	0.59%	(3.96)[d]	586:549	141	1.71%	(4.04)[c]	67:74
AR_{t+1}	1132	0.25%	(1.70)[a]	557:575	141	-1.31%	(-3.11)[c]	55:86
AR_{t+2}	1131	-0.05%	(-0.34)	537:594	141	-0.84%	(-2.00)[b]	61:80
AR_{t+3}	1128	-0.27%	(-1.83)[a]	535:593	141	-0.26%	(-0.62)	64:77
AR_{t+4}	1128	-0.26%	(-1.76)a	511:617	141	0.44%	(1.04)	67:74
AR_{t+5}	1125	-0.10%	(-0.70)	514:611	141	-0.78%	(-1.84)[a]	57:84
$CAR_{t=-5,-2}$	1135	0.83%	(2.79)[c]	588:547	141	-4.03%	(-4.78)[d]	62:79
$CAR_{t=-1,0}$	1135	12.13%	(57)[d]	1069:66	141	21.53%	(36)[d]	131:10
$CAR_{t=+1,+5}$	1135	-0.43%	(-1.30)	537:596	141	-2.76%	(-2.92)[c]	57:84

Note that the numbers in parenthesis are Z-value. The letters a, b, c, and d denote statistical significance at the 0.10, 0.05, 0.01 and 0.001 levels, respectively, using a two-tail test.

For the Good stocks listed on NASDAQ (Not reported in a table, but available upon request), the average abnormal returns one day before the publication, publication day, and the day after publication are, respectively, 17.35, 0.34, and

0.65% (test statistics $Z_{-1} = 82$, $Z_0 = 1.60$, $Z_{+1} = 3.10$).The positive returns at t =0 and t = +1 are not only statistically significant but also high enough to provide institutional and large investors a positive abnormal return net of transaction costs (Mathur and Waheed, 1995). For the NASDAQ good stocks, the average cumulative abnormal return during the period t = -1 to 0 is 17.69% [test statistic Z (-1, 0) = 59].

Table 4 also displays results from the sample of Good Stocks listed at exchanges other than NYSE or NASDAQ. These stocks are listed at exchanges such as OTC, LSE, Frankfurt, TSE, Toronto, TLV, among others. The positive responses are by far greater than those at any other exchanges at t=-1. A significant positive one day return is reported. The cumulative average abnormal return for days -1, 0 is 21.53%, which is the greatest two day return recorded from any sample in this study. In general, the results for the NYSE firms reported in Table 4 appear much more similar to those reported for the overall sample in Table 1 than those for any other exchange stocks. Furthermore, the study found that the Bad Stocks at the NYSE show much more similar patterns to those reported for the overall sample in Table 1 than those stocks at any other exchanges (Not reported in a table, but available upon request).

Summary and Conclusions

The paper examines the information content of the "Money & Investing" section of the Wall Street Journal. More specifically, the paper examines the Journal's daily choice of four "Good stocks" and four "Bad Stocks" from December 3, 2007 to December 31.2009 to determine whether returns on these stocks are abnormal around the publication day.

The Efficient Market Hypothesis suggests that stock prices should respond to a source of relevant secondary information. The daily choice set of four good stocks and four bad stocks serves determined by the WSJ analysts is a source of secondary information readily available to investors. Using the Brown-Warner event study methodology, the study finds statistically significant positive price responses or abnormal returns in Good Stocks on one day after the publication day and statistically significant negative returns in Bad Stocks one day after publication.

In addition, the study found that most of these significant returns are generated by repeated stocks which represent the most noteworthy stocks in terms of multiple appearances in reported news articles. This is partially consistent with the investor overreaction hypothesis: As a result of investors' overreaction to repeated-relevant information about stocks, prices might temporarily rise beyond their underlying fundamental values, remain high for a brief period, and then start falling back toward their intrinsic values.

Second, the significance of abnormal returns prior to the publication day (t = -3, -2, -1) suggests that some investors may be sufficiently well informed to make purchases (sales) of Good (Bad) stocks in advance of publication. Hence, results are consistent with some investors exhibiting informed behavior.

Lastly, the responses of good and bad stocks are asymmetric across exchanges. The positive NASDAQ stock returns at t = +1 are far greater than the positive NYSE stock returns. The positive one day NASDAQ stock returns are high enough to provide institutional and other large investors a positive abnormal return net of transaction costs. For bad stocks, the negative returns of the OTHER exchange stocks are more pronounced than those reported on the NYSE and NASDAQ stocks.

References

Ferreira, Eurico J. and Stanley D. Smith, 1999. "Stock price reactions to recommendations in the *Wall Street Journal* 'Small Stock Focus' column," *Quarterly Review of Economics and Finance*, 29, 379-389.

Conrad, Jennifer, Bardford Cornell and Wayne R. Landsman, 2002. "When is bad news really bad news?" *Journal of Finance*, 57, 2507-2532.

Brown, Stephen J. and Jerold B. Warner, 1985. "Using daily stock returns: the case of event studies," *Journal of Financial Economics*, 14, 3-31.

Mathur, Ike and Amjad Waheed, 1995. "Stock price reactions to securities recommended in *Business Week*'s 'inside *Wall Street*'," *Financial Reviews*, 30, 583-604.

Palmon, Oded, Huey-Lian Sun, and Alex Tang, 1994. "The impact of publication of analysts' recommendations on returns and trading volume," *Financial Reviews*, 29, 395-417.

Yau, Jot, Michael G. Ferri, and Timothy F. Sugrue, 1994. "An analysis of the *Wall Street Journal*'s coverage of corporate news and the research design of event studies," *Journal of Financial Research*, 2, 161-173.

Economic Cycles, Investor Sentiment and the Fundamental Analysis of Value Stocks

Melissa K. Woodley, Steven T. Jones, and James P. Reburn, Samford University

Abstract

We have previously confirmed (Woodley, Jones, and Reburn, 2011) the findings of Piotroski (2000) regarding the ability to use a set of financial statement variables to distinguish future winners from future losers among value stocks, over Piotroski's 1976-1996 test period. However, we also have found that over the following twelve-year period, the results are actually reversed. This paper seeks to tie these disparate findings to the literature on time-varying value premia. Our findings seem to indicate that the ability of the Piotroski (2000) model to pick winners among value stocks improves when economic indicators are average or above-average.

Introduction, Literature Review, and Motivation

Value Stock Returns: Market Efficiency or Inefficiency?

A significant volume of research finds that on average, "value stocks" (i.e., firms with above-average book-to-market ratios) tend to outperform "growth stocks" or "glamour stocks" (i.e., firms with below-average book-to-market ratios). This finding is documented by Rosenberg, Reid, and Lanstein (1984) and, of course, by the classic Fama and French (1992) paper. Further, this "value premium" appears to be robust to a variety of measurement schemes and tests, as documented by Lakonishok, Shleifer, and Vishny (1994), La Porta, Lakonishok, Shleifer, and Vishny (1997), Fama and French (1998), Chan and Lakonishok (2004), and Fama and French (2006).

Indeed, much of the debate appears to center less on the existence of the value premium than on whether the value premium is consistent with the efficient markets hypothesis. Papers that answer this question in the affirmative include the aforementioned Fama and French (1992) paper, along with papers by Penman (1991), Fama and French (1995), and Chen and Zhang (1998). On the other hand, the aforementioned papers by Lakonishok, Shleifer, and Vishny (1994) and by LaPorta et al. (1997) make the case for the view that the value premium demonstrates an inefficiency in the market.

Fundamental Analysis of Value Stocks

Piotroski (2000) points out that several papers prior to his own have looked for ways to identify stocks whose value the market has underestimated as a result of overly pessimistic expectations. In particular, Frankel and Lee (1998), Dechow and Sloan (1997), and LaPorta (1996) have explored this issue.

Piotroski is particularly interested in using financial measures as a means of identifying value stocks that may offer the potential for outsize returns. This is a particularly promising line of research in light of the fact that while value stocks as a group earn above-average returns, a majority of individual value stocks actually produce below-average returns. This result, of course, can only occur if the average outperformance of those value stocks that do outperform the market outweighs the average underperformance of those value stocks that underperform the market; in other words, taken as a group those value stocks that outperform the market apparently do so by a fairly wide margin.

Thus, if one can find a way to isolate those individual value stocks that are likely to provide above-average returns, the overall return advantage that one could obtain by investing in value stocks can, at least in theory, be magnified considerably. Holthausen and Larcker (1992), Lev and Thiagarajan (1993), and Abarbanell and Bushee (1998) have all made strides in using fundamental analysis to estimate future market performance.

The Piotroski Methodology

Since our work was initially motivated by an attempt to investigate the replicability of Piotroski's (2000) results, our prior paper (Woodley, Jones, and Reburn, 2011) devotes a great deal of attention to a description of his methodology. That description is presented in a somewhat shorter form below.

For each of the years in his 1976-1996 window, Piotroski (2000) sorts the universe of firms on which stock returns data is available into quintiles, based on the various firms' book-to-market ratios. Those stocks in the top quintile are identified as "value stocks" for purposes of his study. He separately sorts the universe of stocks into market value terciles.

Each firm in the value stock sample is then evaluated on nine separate factors, with a score of 1 or 0 assigned for each factor. Scores of 1 are assigned for the following factors related to profitability: a positive return on beginning-of-year assets, ignoring extraordinary items (ROA); a positive measure for cash flow from operations as a percentage of beginning-of-year assets (CFO); a positive year-over-year change in ROA (ΔROA); and a negative difference between ROA and CFO (ACCRUAL). Scores of 1 are likewise assigned for the following factors related to leverage, liquidity, and source of funds: a year-over-year decline in long-term debt as a percentage of average total assets (ΔLEVER); a year-over-year increase in the current ratio (ΔLIQUID); and the absence of an issuance of common equity within the past year (EQ_OFFER). Finally, scores of 1 are assigned for the following factors related to operating efficiency: a positive year-over-year change in the gross profit margin (ΔMARGIN); and a positive year-over-year change in the ratio of sales to beginning-of-year assets (ΔTURN). Scores of 0 are assigned in each case where a given test is not met.

The sum of each firm's scores on these factors is identified as the firm's "F_Score." Mean and median returns are calculated for each score from 0 through 9, inclusive. Separate calculations of mean and median returns are made for the pooled set of firms with scores of 0 and 1 (identified as "Low Score" firms) and for the pooled set of firms with scores of 8 and 9 (identified as "High Score" firms). Results for "High Score" firms are compared to those of "Low Score" firms, and to those for the set of value stocks as a group.

Piotroski's Findings, and Ours

As shown in his Table 3, Piotroski (2000) finds a fairly consistent pattern in which market-adjusted returns increase as the F_Score increases. Further, this superiority of returns for "High Score" firms is highly statistically significant across a number of comparisons. In his Table 4, Piotroski shows that the outperformance of "High Score" firms is extremely strong among small firms, is also quite impressive among mid-size firms, and is rather weak among large firms.

Our own results (Woodley, Jones, and Reburn, 2011) for the Piotroski (2000) test period of 1976-1996 are qualitatively similar to Piotroski's, confirming that "High Score" firms tend to produce above-average returns. However, during a subsequent test period from 1997-2008, our findings are exactly reversed from those during the Piotroski test period. On average, "High Score" firms underperform value stocks as a group by a wide margin, and underperform "Low Score" firms by an even wider margin. Further, we find that both our agreement with Piotroski's findings that "High Score" firms outperform over his test period, and our result that "High Score" firms underperform over the ensuing 11-year period, are largely invariant to firm size.

Motivation: Tying Disparate Findings Over Different Time Periods to Time-Varying Premia

Our findings in Woodley, Reburn and Jones (2011) lead to a question: why is it that the very factors that lead to outperformance during the Piotroski (2000) test period of 1976-1996 actually seem to predict underperformance during the ensuing 12-year period?

This paper seeks to tie these seemingly conflicting findings to the literature on time-varying value premia by analyzing the ability of Piotroski's fundamental value stock screen to identify winners and losers conditional on expected business conditions and investor sentiment. Because value stock returns are more sensitive to shifts in investor sentiment (Baker and Wurgler, 2006) and market cash-flow shocks (Campbell, Polk, and Vuolteenaho, 2010), we hypothesize that the ability of value stock accounting screens to identify potential winners and losers will be improved by conditioning observed returns on forecasts of economic conditions.

Data and Methodology

As described in our previous work on this topic (Woodley, Jones, and Reburn, 2011), we employ Compustat data regarding the numbers from various firms' financial statements, and CRSP data regarding both market returns and market capitalization. Our methodology is described in detail in that paper, and those portions of the methodology that are replicated in this paper are briefly summarized below.

For any given fiscal year from 1976 through 2008, inclusive, a given firm's book-to-market ratio is calculated as of the ending date of the prior fiscal year. (Thus, market-to-book ratios are calculated going back to fiscal year 1975.) The various book-to-market ratio observations are sorted into quintiles. Subject to completeness of data availability, any observation that

falls into the top book-to-market quintile is included in the universe of value firms. The aforementioned financial statement factors are calculated for the fiscal year in question, and market-adjusted returns are calculated for the year beginning in the fifth month after the end of that fiscal year. Thus, if a firm uses a calendar year, then for purposes of this study its returns for a given fiscal year are the returns for the period beginning on May 1 of the following year, and ending on April 30 of the year after that. As a result, an "ending date" of 2008 for our test period actually corresponds to returns data that, in the cases of some firms, extends as far as April 30, 2010. As of the time the tests described in this paper were run, returns data were not yet available for 2011.

While Tables 1 and 2 from Woodley, Jones, and Reburn (2011) are replicated in the current working paper, for purposes of space we have instead summarized the results below for this Proceedings version of the paper. Table 1 from Woodley, Jones and Reburn (2011) provides basic statistical information regarding the financial characteristics of the subset of value stocks as compared to the entire universe of stocks in the sample. On three of the measures of profitability – ROA, DROA, and CFO – the subset of value stocks has a substantially smaller proportion of firms showing a positive signal than does the entire set of firms. This holds true regardless of whether we are considering the entire period, the 1976-1996 sub-period, or the 1997-2008 sub-period. (The opposite holds true with the ACCRUAL measure, although the differences appear minimal.) It also is worth noting, however, that for both the entire sample and the value stock sub-sample, there is a double-digit decline over time in the percentage of stocks showing a positive signal on the ROA measure; this may be viewed as providing some evidence of a decline in profitability in general.

Regarding the measures of leverage and liquidity, the subset of value stocks consistently shows a noticeably (albeit modestly) smaller proportion of firms with a positive signal on the DLIQUID measure than does the overall sample. On the DLEVER measure, while the difference is in the same direction, the degree of this difference is quite small.

Finally, in the area of operating efficiency, on both the DMARGIN and DTURN measures the subset of value stocks consistently produces a smaller proportion of positive signals than does the overall sample. As with the aforementioned difference in the DLIQUID measure, these differences are relatively modest, but noticeable.

Table 2 from Woodley, Jones and Reburn (2011) compares the results of a value strategy versus those of a growth strategy. Over our entire 33-year test period (from 1976-2008, inclusive), both mean and median returns are higher (with a p-value < 0.0001) for a value strategy. This finding also holds for both sub-periods. Further, on a year-by-year basis the number of years in which returns are significantly higher (at the 10% level) for a value strategy than for a growth strategy dwarfs the number of years in which the reverse holds true. The ratio is nearly four-to-one in the case of mean returns, and is exactly three-to-one in the case of median returns.

In this paper, we perform the following additional tests. To address the issue of time-varying investment premia, we perform two separate tercile sorts. In the first such sort, we obtain data regarding the Index of Leading Economic Indicators from the website of the St. Louis Federal Reserve. Leading Indicator readings over the available dates (which begin in January 1982, several years after the start of the Piotroski (2000) test period) are sorted into terciles. Firm observations are placed into the appropriate category – "Low Growth," "Medium Growth," or "High Growth" – depending on the leading indicator tercile into which the portfolio formation date falls.

Then, a series of comparisons are run. We compare the subset of "High F_Score" firms to both the subset of "Low F_Score" firms, and the set of value stocks as a whole. Each of these comparisons is made, based on both mean and median returns, over the Piotroski (2000) test period (shortened somewhat by the aforementioned data availability limitation for the leading indicators), over our subsequent test period of 1997-2008, and over the combined test period. These comparisons are made within the "Low Growth," Medium Growth," and "High Growth" subsets.

In the second tercile sort, we obtain data regarding the University of Michigan Consumer Sentiment Index from the website of the St. Louis Federal Reserve. With this data series, available dates go back to January 1978. These readings are sorted into terciles. Firm observations are placed into "Low Sentiment," "Medium Sentiment," and "High Sentiment" categories, depending on the consumer sentiment tercile into which the portfolio formation date falls. Comparisons are then run, analogous to the comparisons described above for the Leading Indicator terciles.

Results

Relative Returns as They Relate to Leading Economic Indicators

Tables 3, 4, and 5 from our current working paper display the results obtained when the sample is subdivided into those data points that fall into the low, medium, and high growth indicator terciles of the overall test period. For each tercile tested within each time period, there are four tests of statistical significance, as the "High Score" subset is compared both to the "Low Score" subset and to the overall set of value firms, on both the mean and median return measures. Although the tables

themselves are omitted from this Proceedings version of the paper for space purposes, the contents of these tables are described below.

In Table 3, we focus specifically on the Piotroski test period, with the aforementioned limitation that our data prevents us from going back further than January of 1982. Nonetheless, this still affords a 15-year test period (1982-1996, inclusive), which comprises a large majority of the overall Piotroski (2000) test period of 1976-1996.

The first noteworthy item regarding this table is that within this sub-period, over 47% of the observations (5,512 observations out of 11,627) fall within the high-growth tercile of the overall test period. When leading indicators are poor, the "High Score" firms statistically outperform at the 10% level or better in two of the four tests (specifically, in the comparisons of both mean and median returns for "High Score" versus "Low Score" firms), and very nearly in a third test (median returns for "High Score" firms vs. all firms). Both when the leading indicators forecast medium growth and when they forecast high growth, the "High Score" firms statistically outperform at the 10% level or better in three of the four tests. Further, in five of these six cases the p-value is well below 0.01, and in the one remaining comparison the p-value is 0.0117. On balance, the statistical significance of the outperformance of the "High Score" firms is arguably somewhat stronger in medium-growth scenarios than in high-growth scenarios; but, is clearly stronger in both medium-growth and high-growth scenarios than in low-growth scenarios.

Table 4 displays the results of analogous tests over the period subsequent to that tested by Piotroski (2000). Leading indicators tended to be much weaker on average during this period than during the prior period. In fact, nearly 48% of the observations (4,682 out of 9,767 observations) fall within the low-growth tercile of the overall test period.

During this sub-period, when leading indicators are poor the "High Score" firms significantly underperform in two of the four tests – specifically, in the two tests for mean returns. Since the "High Score" firms significantly underperform on the tests for equality of means, while non-significantly outperforming on the tests for equality of medians, it seems probable that there are a number of strongly underperforming "High Score" firms that are driving the underperformance of the mean returns.

When leading indicators forecast medium growth, the "High Score" firms significantly underperform in one of the four tests – specifically, the comparison of mean returns for the "High Score" firms versus all firms, with a p-value of well below 1%. In a second case – median returns for this same comparison – the underperformance of the "High Score" firms has a p-value of approximately 12.5%. Mean and median returns are also lower for "High Score" firms than for "Low Score" firms, but neither of these results is statistically significant. Finally, when leading indicators forecast high growth, the "High Score" firms significantly underperform in one of the four tests. As with the medium growth breakdown, this significant result is found when comparing the mean returns of "High Score" firms versus those of all firms; the p-value is 0.0194.

Thus, a comparison of the two periods tells us that the outperformance of "High Score" firms in the Piotroski (2000) test period tends to be weakest when economic indicators are weak. Meanwhile, the underperformance of "High Score" firms in the subsequent period tends to be a bit more pronounced when economic indicators are weak. Taking these two findings combined, it would seem that the Piotroski (2000) model is at least somewhat more likely to pick winners, and less likely to pick losers, during periods when economic indicators are favorable than during periods when economic indicators are unfavorable.

Table 5, which deals with the combined test period, seems to confirm this result. Recall that this overall period combines an earlier sub-period in which "High Score" firms tend to outperform with a later sub-period in which "High Score" firms tend to underperform. During this overall period, when we focus on observations that fall within the low-growth tercile we find that the "High Score" firms show no statistically significant differences in any of the four tests for equality. (Despite the lack of significance of these differences, it may be worth noting that regardless of whether we are comparing "High Score" firms to "Low Score" firms or to the entire set of value stocks, the "High Score" firms show lower mean returns but higher median returns. This, again, would tend to indicate the presence of some negative outliers among the "High Score" firms.)

When we look instead at observations that fall within the medium-growth tercile, we find that the coefficient on the differences is now positive in three of four cases. One of these positive differences (specifically the difference in median returns of "High Score" firms versus those of "Low Score" firms) is significant, with a p-value of just below 0.01.

Finally, when we analyze those observations that fall within the high-growth tercile, we find that the coefficient on the differences is now positive in all four cases, and that three of these four differences are statistically significant. Both differences in median returns are easily significant at the 1% level. The difference in mean returns, when comparing the "High Score" firms versus the entire sample, has a p-value of 0.0128, easily significant at the 5% level but not quite significant at the 1% level. The only non-significant result—the difference in the mean returns of "High Score" versus "Low Score" firms – has a p-value of 0.1218.

Thus, for the overall test period the overall trend seems to be that a strategy of picking "High Score" firms works better during periods when economic indicators point to high growth.

Consumer Sentiment

Tables 6, 7, and 8 from our current working paper perform a series of tests similar to those in Tables 3, 4, and 5, using consumer confidence rather than economic indicators as the basis for subdividing the sample into terciles. Oddly enough, while leading indicators tended to be stronger during the Piotroski (2000) test period than during the subsequent period, consumer sentiment readings actually tended to be weaker during the Piotroski (2000) test period. Once again, the tables themselves are omitted from this Proceedings version of the paper due to space considerations, but are described in some detail below.

Table 6 looks at the Piotroski (2000) test period. As with the leading indicators examined in Table 3, our data do not allow us to go back to the very beginning of the Piotroski test period. However, they do allow us to go back to January of 1978; so, the vast majority of the Piotroski test period is included. Over this sub-period, when consumer sentiment is low the "High Score" firms produce higher mean and median returns than do either the "Low Score" firms or the overall sample. In two of these four tests, the difference in returns is significant at the 10% level. (This occurs in the "High Score versus All" comparison of median returns, and in the "High Score versus Low Score" comparison of mean returns.)

Results are much stronger when consumer sentiment is classified as medium or high. Within both the medium and high consumer sentiment terciles, median returns are consistently higher for "High Score" firms at the 1% level, regardless of whether the comparison is to "Low Score" firms or to the entire sample. Mean returns are significantly higher for "High Score" firms at the 10% level in all four comparisons, and at the 5% level in two of these four comparisons. (Specifically, the difference in mean returns is significant at the 5% level during times of medium sentiment when the "High Score" firms are compared to the "Low Score" firms, and during times of high sentiment when the "High Score" firms are compared to the overall set of value stocks.)

Thus, during the Piotroski test period both mean and median returns are higher for "High Score" firms; but, these differences are far more significant during periods of medium and high sentiment than during periods of low sentiment.

Table 7 examines at the post-Piotroski (2000) test period. When consumer sentiment is low, "High Score" firms have higher mean returns and lower median returns, both in comparison to the "Low Score" firms and in comparison to the overall sample. However, while the measured differences on both measures are large, none of these results are statistically significant. (There are only 21 "High Score" observations in this subsample, thus weakening statistical power.)

During medium-sentiment periods, "High Score" firms consistently show higher mean and median returns, but these results are significant in only one of the four tests that we perform. (Specifically, median returns are significantly higher, with a p-value of 0.0113, for "High Score" firms than for "Low Score" firms.)

On the other hand, during periods of high consumer sentiment, the "High Score" firms produce consistently below-average returns. In the case of the mean returns, the p-value of this result is less than 0.0001, regardless of whether the comparison is to the "Low Score" firms or to the overall sample. In the case of the median returns, the underperformance of the "High Score" firms relative to the overall sample has a p-value of 0.0249, while the underperformance of the "High Score" firms relative to the "Low Score" firms is not statistically significant.

Results during the combined period seem less straightforward than do the results during the two sub-periods. During times of low sentiment, mean returns of "High Score" firms are below both those of "Low Score" firms and those of the overall set of value firms, with both differences easily significant at the 1% level. However, median returns of "High Score" firms are significantly above those of the overall set of value firms (p-value of 0.0168). Median returns of "High Score" firms are below those of "Low Score" firms, but the difference is not significant.

During periods of medium sentiment, "High Score" firms produce significantly higher median returns than do "Low Score" firms, with a p-value of less than 0.0001. While "High Score" firms also produce higher median returns than do value stocks as a group, and higher mean returns than both "Low Score" firms and the set of all value stocks, none of these results is statistically significant.

Finally, during times of high sentiment, "High Score" firms produce median returns that are significantly higher than those of both "Low Score" firms and the overalls set of value stocks. However, "High Score" firms produce mean returns that lag those of both "Low Score" firms and the overall set of value stocks; further, in the case of the comparison to the overall set of value stocks, the result is easily significant at the 10% level, and nearly reaches the 5% significance level.

In Tables 6, 7, and 8, there is again some fairly strong evidence of negative outliers in the "High Score" group. With few exceptions, the "High Score" firms' relative performance looks better when the basis of comparison is the median return than when it is the mean return.

Conclusion and Future Research Direction

In previous work (Woodley, Jones, and Reburn, 2011), we have confirmed the findings of Piotroski (2000) regarding the ability to use strong performance on a set of financial statement variables to distinguish future winners from future losers among value stocks, over the time period that he tested. However, in that same study we also have found that over twelve-year period following Piotroski's test period, the results are actually reversed: firms that have high book-to-market ratios and that have high scores on Piotroski's "F_Score" measure actually produce returns below those of other value stocks. Both of these sets of results continue to hold when we adjust for firm size.

This paper seeks to tie this set of findings to the literature on time-varying value premia by analyzing the ability of Piotroski's fundamental value stock screen to identify winners and losers conditional on expected business conditions and investor sentiment. Because value stock returns are more sensitive to shifts in investor sentiment (Baker and Wurgler, 2006) and market cash-flow shocks (Campbell, Polk, and Vuolteenaho, 2010), we hypothesize that the ability of value stock accounting screens to identify potential winners and losers will be improved by conditioning observed returns on forecasts of economic conditions.

Probably the most fascinating finding in this study has to do with the difference between our results when we break down the sample based on economic indicators, as opposed to when we break down the sample based on consumer sentiment. The results do seem to indicate that the ability of the Piotroski (2000) model to pick winners among high book-to-market stocks improves when economic indicators are average or above-average, as opposed to below-average. This general result holds, for the most part, for the overall period. Further, outperformance of "High Score" stocks is relatively more likely to be statistically significant during the earlier period, and underperformance of "High Score" stocks is relatively less likely to be statistically significant during the later period, when economic indicators are average or above-average.

And, at least during the Piotroski test period, there seems to be at least some indication that an analogous pattern holds with regard to consumer sentiment: by some measures, at least, the Piotroski model performs better during times of medium or high sentiment than during times of low sentiment. However, during the later sub-period, the pattern seems to reverse somewhat: For the most part, the relative performance of "High Score" firms actually appears to weaken as consumer sentiment improves.

Stated another way, there appears to be greater support for the notion that the Piotroski (2000) investment strategy works best at those times when actual economic statistics provide grounds for optimism about the near-term future of the economy, than for the notion that this strategy works best at those times when consumers "feel" optimistic.

Another interesting set of findings – which may or may not be related to the set of findings described above – is that is that the earlier test period, in which the Piotroski (2000) strategy seems to work rather well, contains a disproportionate number of data points occurring during times of strong economic indicators, but also contains a disproportionate number of data points occurring during times with low consumer sentiment. The later test period, in which the Piotroski strategy seems not to work well, contains more data points during times of weak economic indicators, and more data points during times with high consumer sentiment.

With regard to the leading indicators data in particular, it would be tempting to interpret this set of results as indicating that the previously-observed strong results for "High Score" firms during the earlier period, and weak results during the later period, are merely proxies for the fact that better results seem to occur during times when leading indicators are average or above-average. However, that would almost certainly be an over-interpretation of our findings; for instance, within the earlier period, "High Score" firms produce higher average returns than do other value stocks during periods of strong, average, and weak economic indicators alike. Further, during the later period, regardless of whether we are looking at the strong-, average-, or weak-indicator sub-sample, the only statistically significant differences between "High Score" firms and other value stocks are those in which the "High Score" firms under-perform.

So, while it is possible that the Piotroski (2000) strategy's relative strength during the earlier period, and its relatively weakness during the latter period, can be explained by the relatively large number of strong-indicator observations during the former and the relatively large number of weak-indicator observations during the latter, there is obviously more to the story than that. For whatever reason or reasons, the Piotroski strategy worked better during the earlier period than during the latter; at most, this overall pattern was merely exacerbated by the fact that the earlier period contained more observations taken during times when economic indicators were good.

Based on all of the above observations, we believe that a fruitful area for further testing would be to compare the relative performance of "High Score" stocks based on various combinations of consumer sentiment and leading indicators. In other words, do "High Score" stocks produce their best returns when consumer sentiment is more pessimistic than leading indicators would suggest that it should be, and produce their worst returns when consumer sentiment is more optimistic than leading indicators would suggest that it should be? If so, this might explain why Piotroski's (2000) model worked better

during a period disproportionately characterized by relatively strong economic indicators and relatively weak consumer sentiment, than during a period in which the opposite characteristics seemed to hold.

References

Abarbanell, Jeffery S., and Brian J. Bushee. Spring 1997. "Fundamental Analysis, Future Earnings, and Stock Prices." *Journal of Accounting Research* 35(1): 1-24.

Ang, Andrew, and Joseph Chen. January 2007. "CAPM Over the Long Run: 1926-2001." *Journal of Empirical Finance* 14(1): 1-40.

Baker, Malcom and Jeffrey Wurgler. August 2006. "Investor Sentiment and the Cross-Section of Stock Returns." *Journal of Finance* 61(4): 1645-1680.

Campbell, John Y., Christopher Polk, and Tuomo Vuolteenaho. January 2010. "Growth or Glamour? Fundamentals and Systematic Risk in Stock Returns." *Review of Financial Studies* 23(1): 305-344.

Chan, Louis K.C., and Josef Lakonishok. January/February 2004. "Value and Growth Investing: Review and Update." *Financial Analysts Journal* 60(1): 71-86.

Chen, Nai-fu, and Feng Zhang. October 1998. "Risk and Return of Value Stocks." *Journal of Business* 71(4): 501-35.

Dechow, Patricia M., and Richard G. Sloan. January 1997. "Returns to Contrarian Investment Strategies: Tests of Naïve Expectations Hypotheses." *Journal of Financial Economics* 43(1): 3-27.

Fama, Eugene F., and Kenneth R. French. June 1992. "The Cross-Section of Expected Stock Returns." *Journal of Finance* 47(2): 427-65.

Fama, Eugene F., and Kenneth R. French. March 1995. "Size and Book-to-Market Factors in Earnings and Returns." *Journal of Finance* 50(1): 131-55.

Fama, Eugene F., and Kenneth R. French. December 1998. "Value versus Growth: The International Evidence." *Journal of Finance* 53(6): 1975-1998.

Fama, Eugene F., and Kenneth R. French. October 2006. "The Value Premium and the CAPM." *Journal of Finance* 61(5): 2163-2185.

Frankel, Richard, and Charles M.C. Lee. June 1998. "Accounting Valuation, Market Expectation, and Cross-Sectional Stock Returns." *Journal of Accounting and Economics* 25(3): 283-319.

Holthausen, Robert W., and David F. Larcker. June-September 1992. "The Prediction of Stock Returns Using Financial Statement Information." *Journal of Accounting and Economics* 15(2-3): 373-411.

Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny. December 1994. "Contrarian Investment, Extrapolation, and Risk." *Journal of Finance* 49(5): 1541-78.

LaPorta, Rafael. December 1996. "Expectations and the Cross-Section of Stock Returns." *Journal of Finance* 51(5): 1715-42.

LaPorta, Rafael, Josef Lakonishok, Andrei Shleifer, and Robert Vishny. June 1997. "Good News for Value Stocks: Further Evidence on Market Efficiency." *Journal of Finance* 52(2): 859-74.

Lev, Baruch, and S. Ramu Thiagarajan. Autumn 1993. "Fundamental Information Analysis." *Journal of Accounting Research* 31(2): 190-215.

Lintner, John. February 1965. "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets." *Review of Economics and Statistics* 47(1): 13-37.

Loughran, Tim. September 1997. "Book-to-Market across Firm Size, Exchange, and Seasonality: Is There an Effect?" *Journal of Financial and Quantitative Analysis* 32(3): 249-268.

Penman, Stephen H. Spring 1991. "An Evaluation of Accounting Rate-of-Return." *Journal of Accounting, Auditing, and Finance* 6(2): 233-55.

Piotroski, Joseph D. 2000. "Value Investing: The Use of Historical Financial Statement Information to Separate Winners from Losers." *Journal of Accounting Research* 38(Supplement): 1-41.

Rosenberg, Barr, Kenneth Reid, and Ronald Lanstein. Spring 1985. "Persuasive Evidence of Market Inefficiency." *Journal of Portfolio Management* 11(3): 9-16.

Sharpe, William F. September 1964. "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk." *Journal of Finance* 19(3): 425-42.

Woodley, Melissa K., Steven T. Jones, and James P. Reburn. December 2011. "Value Stocks and Accounting Screens: Has a Good Rule Gone Bad?" *Journal of Accounting and Finance* 11(4): 87-104.

Using Currency ETFs to Hedge Foreign Exchange Risk

Robert B. Burney, Coastal Carolina University

Abstract

The emergence of currency exchange traded funds (CETFs) presents an alternative vehicle for both speculation and hedging in the currency markets. In particular, because currencies ETFs typically have small minimum transaction sizes and trade like equities, they represent an interesting opportunity for the small investor or small business operator.

Overview

The trade press has presented various practitioner oriented articles concerning the usefulness of currency ETFs in hedging foreign currency risk. However, such discussions typically do not approach the topic systematically. Generally, such treatments fail to recognize the similarities and differences between ETF based hedges and other hedging techniques using futures contracts, options contracts, or money market transactions. Most notably, many of the practitioner oriented articles discount the use of currency ETFs in hedging applications because of the necessity of tying up capital during the hedge. This criticism lacks merit since other well established techniques for dealing with foreign exchange risk also tie up capital during the hedging period. In following sections this paper reviews the issues which arise in using currency ETFs in managing foreign exchange risk. Transaction details specific to currency ETFs are then discussed. Examples are given contrasting various hedging techniques, with special emphasis on incorporating the cost of capital in the analysis. Finally, recommendations are made for implementation of various types of currency ETF based hedges.

Management Issue

Table One: Example Derivative Contract Sizes

PHLX:	
Euro Options	10,000 euros
CME:	
Euro Futures*	125,000 euros
E-mini Futures	62,5000 euros
E-micro Futures	12,5000 euros

*CME Options are limited to larger contracts.

The basic issue at hand is the management of foreign currency risk for those market participants who face only very small or occasional foreign currency exposures. For such market participants, existing liquid market derivative securities simply do not match the transaction scale. The table to the left summarizes this situation for the dominant U.S. market derivatives for the euro. In this case, the smallest notional principal amount involves 10,000 euros. A smaller size transaction would force the market participant to, in effect, take on a residual exposure of opposite nature to the initial exposure.

While services for small scale market participants do exist among the retail foreign exchange dealers, these arrangements have a reputation of being of a disadvantageous cost structure with account details which may create more difficulties for the potential hedger (full margin calls, etc.). Also, the smaller FX dealer based derivative contracts or minor electronic exchanges suffer from illiquidity which may negatively impact pricing.

Currency ETFs

In recent years numerous currency ETFs have been introduced. These include ETFs which cover most of the major currencies. Variations include both long and short position ETFs, and more recently, double and triple long and short varieties. Table Two presents a sample listing of available currency ETFs as discussed on the *Artremis.com* website.

Currency ETFs (CETFs) strive to match the change of the target currency in the specified proportion on a day -to-day basis. Essentially, the CETF substitutes for a long or short position in the target currency. Because the CETFs are traded on a share basis, an investor can take any conceivable position depending on the number of shares purchased. Based on varying notional amounts, most attempt to capture the percentage change of the target currency. While most applications of both speculation and hedging in foreign currencies can be accomplished with outright currency trades and traditional derivatives, the ease of trading CETFs is attractive to the new or occasionally foreign currency impacted investor. There is no set denomination per share of such CETFs. Some CETFs are quoted in multiples of the underlying currency, while others

are based on an arbitrary notional principal. Recent per share values range from $15.87 to $132.72 for the CETFs presented in the table.

Table Two: Examples of Currency ETFs*

Symbol	Name	Fund Family	Currency
FXE	CurrencyShares Euro Trust	Rydex	Euro
EU	WisdomTree Dreyfus Euro	WisdomTree	Euro
ERO	iPath EUR/USD Exchange Rate ETN	iPath	Euro/U.S. dollar exchange rate
ULE	Ultra Euro ProShares	ProShares	2x EUR/USD daily price change
EUO	UltraShort Euro ProShares	ProShares	2x inverse EUR/USD daily price change
URR	Market Vectors Double Long Euro	Market Vectors	2x long euro
DRR	Market Vectors Double Short Euro	Market Vectors	2x short euro
UUP	PowerShares DB US Dollar Index Bullish	PowerShares	US Dollar
UDN	PowerShares DB US Dollar Index Bearish	PowerShares	short US Dollar
FXY	CurrencyShares Japanese Yen Trust	Rydex	Japanese Yen
JYF	WisdomTree Dreyfus Japanese Yen	WisdomTree	Japanese Yen
JYN	iPath JPY/USD Exchange Rate ETN	iPath	Japanese Yen/U.S. dollar exchange rate
YCL	Ultra Yen ProShares	ProShares	2x JPY/USD daily price change
YCS	UltraShort Yen	ProShares	2x short , JPY/USD daily price change
FXC	CurrencyShares Canadian Dollar Trust	Rydex	Canadian Dollar

* Source: Artimis.com

The minimum number of shares which an individual investor could trade depends on the individual brokerage housing the account. This conceivably could be a few as one share, and should not be confused with the inter-institutional "creation units" of much larger magnitude. Commissions on CETF trades are also subject to wide variation, with typical commissions at discount brokerages below $10 per trade. Some brokerages also offer commission-free trading on select EFTs.

The basic CEFTs involve foreign currency denominated bank accounts. The inverse or leveraged CETFs also use derivatives to attempt to meet their objectives. (Lachini, 2011)

Examples of Managing Foreign Exchange Risk with CETFs

Foreign currency denominated equity index fund investment

Suppose a U.S. based investor wished to invest in a euro denominated equity index fund such as EZU. By investing in such a fund, the investor is simultaneously exposed to gains or losses on the Eurozone equity market and gains or losses from changes in the value of the euro relative to the dollar. If the investor wishes to speculate on the value of the euro along with his equity investment, so much the better. But, if the investor is interested only in the equity returns, he or she has taken on unwanted currency risk.

One approach to cancel out the currency risk would be to take the desired long position in the EZU and the take an equal position in a short euro ETF such as EUO. Any positive or negative currency effect in the dollar value of the EZU position would be offset by the EUO position leaving only the euro denominated index gain or loss.

Of course, this approach ties up capital in the EUO position, but using double short CETFs like the EUO would require only half the capital assuming the CETF does indeed track consistently at -200% of the actual currency change. The cost of the hedge would be the opportunity cost on the funds in the CETFs, not unlike a traditional money market hedge.

Figure One: Hedging a foreign Currency Cash Flow using Forward Contract

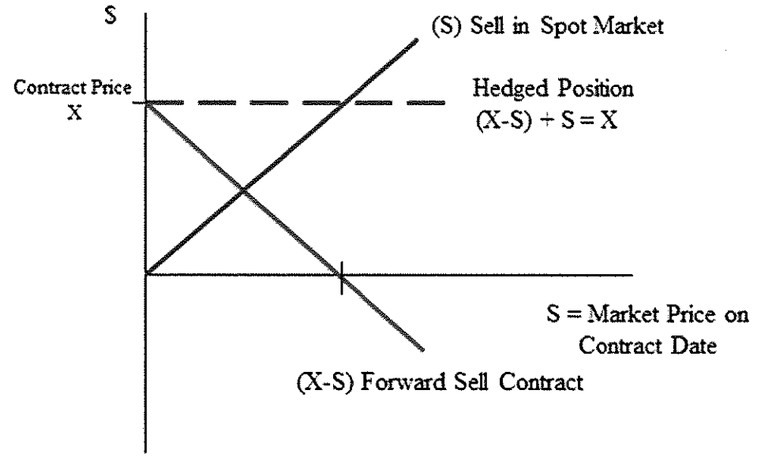
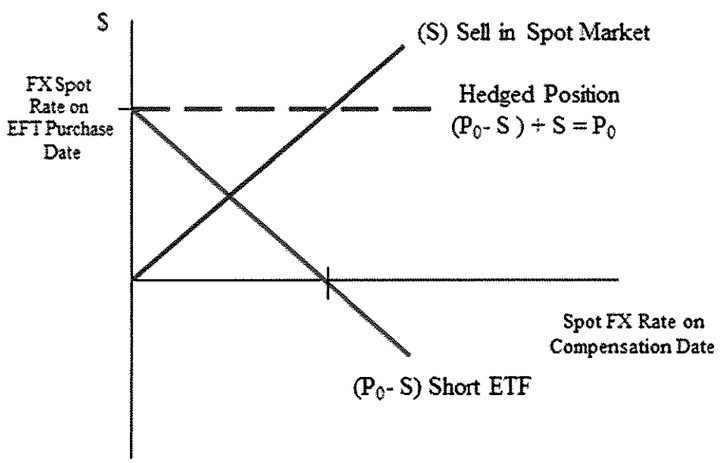


Figure Two: Hedging a foreign Currency Cash Flow using Currency ETF



In understanding the hedge positions, it must be noted that taking the position in the CETF essentially creates the gain or loss equivalent of forward market position with a contract price equal to the spot price on the day the CETF is purchased and a (potentially) infinitely adjustable contract date. The difference relative to forward contracts is that capital is actually being tied up, so this hedge position is not pure in the forward contract sense. Consequently, we must account for the opportunity cost of capital as is done for a money market hedge.

However, in the non-leveraged CETFs, it is generally the case that the fund's holdings are in the held in foreign currency denominated bank accounts earning a money market rate of interest or above. In this instance, the cost of capital tied up in the hedge is already covered.

Small Scale Foreign currency receivable

In many cases of occasional small scale international transactions, the minimum contract size for traditional derivatives may not be met. Suppose an American faculty member takes a short-term summer teaching position at a German university receiving net compensation of 5,000. Due to the size of the cash flow, this amount cannot be easily matched with standard exchange traded futures or options contracts due to contract size limits or lack of participant knowledge or existing derivative trading account.

In this case, the faculty member could easily create a hedge on the dollar equivalent of the compensation by buying a nearly matching amount of shares in a short euro CETF. Any loss in the value of the euro compensation would be offset by the gain on the CETF. The CETF position is held open until the day the euro denominated compensation were received. Again, the cost of the hedge would be the opportunity cost on the funds invested in the ETF. And, depending on the particular CETF used, the cost of funds would be at least partially offset by the rate earned on the foreign bank account holdings of the CETF. A possible complication is that in some instances, the funds interest earnings would be in foreign currency borrowing, so the analyst would need to account for this in estimating the opportunity cost rate of interest.

Issues for Further Research

While the basic construct seems robust, justification for use of the technique depends on some operational details. First, while the CETF management objectives are clearly stated (ex. track double the inverse of the change in the underlying currency), the effectiveness of the CETF managers should be tested. For example, a double long CETF should correlate highly to 200% of the underlying currency s value change. If not, the utility of using double or triple long or short CETFs is diminished. In some instances, operationalization of fund objectives could lead to unintended or unexpected consequences. As example, one should account for the oft cited "constant leverage trap" in leveraged ETFs (Yates, 2007).

Second, while many widely used currency risk hedging techniques (money market hedges, back-to-back loans) do tie up capital during the duration of the hedge, the proper opportunity cost treatment for CETF based hedging must be more completely developed. Surely, those who dismiss CETF based hedging due to this factor overstate the severity of the issue. Nonetheless, a systematic incorporation of capital availability and costs must be established.

Summary

This paper has explored the use of Currency Exchange Traded Funds in hedging foreign exchange exposure of small or occasional market participants. While CETFs appear to hold potential for market participants of this category, more precise development of implementation criteria is needed.

References

Cain, Alexandra. 2011. Small Exporters Getting Savvy About Currency Risk. ***Sydney Morning News***, July 27.
Currency Exchange Traded Funds (ETFs). 2011. *Artremis.com*. September 14.
Goodboy, David. 2008. FOREX spot trades vs. Currency Futures." *TradingMarkets.com*. May 21.
Hudacheck, Dennis. 2011. Currency-Hedged ETFs Lower Volatility." ***Index Universe***. November 16.
Lachini, Michael. 2011. "Leveraged and Inverse ETFs: Not Right for Everyone." Schwab.com. October 20.
Lachini, Michael and Tatjana Michel. 2011. "Currency ETFs: The Facts and Current Dollar Outlook." Schwab.com. October 1.
Hedge Against Exchange Rate Risk With Currency ETFs." 2009. ***Investopedia***. August 12.
Yates, Tristan. 2007. "The Case Against Leveraged ETFs." Seeking Alpha.com. May 17.

How Does CEO Career Origin Influence Firm's Risk-Taking?

Candra Chahyadi, Eastern Illinois University
Pamitra Wineka, University of Illinois at Urbana-Champaign

Abstract

We examine how a CEO career origin affects firm risk-taking and we find that outsider CEOs make more risky investment decisions than insider CEOs. After controlling for compensation sensitivity variables, we find outsider CEOs are still significantly more risk-taking and the board of directors may hire outsider CEOs because of the risk-taking reason. We also find that compensation structure difference is not the reason why outsider CEOs take more risk than insider CEOs. Interestingly outsider CEOs increase the risk of the firm by investing more in R&D and less in capital expenditure, but *not through* increased leverage.

Introduction

How much does a CEO career origin matter for corporate risk-taking? Is the level of corporate risk-taking determined by firm characteristics (such as firm size or past stock returns) or by CEO personal/compensation characteristics or by the combination of the two types of aforementioned characteristics? If risk-taking is somehow influenced by CEO characteristics, can firms incentivize their CEOs to take more risks? How much financial incentives do risk-taking CEOs need to take more risks? These are some important questions whose answers are not easily available because the causal relation between CEO and firm risk-taking is complex and ambiguous. Finding answers to these questions is crucial because level of risk-taking, to some extent, positively affects the value of the firm and by understanding the causal relation (or lack thereof) between CEO characteristics and firm risk-taking, we can find more efficient and cost-effective ways to maximize firm value. The novel contribution of this paper is to establish the relation between a CEO career origin and corporate risk-taking and to explain why firms are now more likely to hire an outsider candidate to fill their vacant CEO positions than before.

One possible way that a CEO is related with firm risk-taking is through CEO compensation, especially through the equity-based compensation. Equity-based compensation is an important mechanism that firms use to align interests of managers and shareholders. Bebchuk and Grinstein (2005) document the increasing occurrence of equity-based compensation practice since 1993. The proliferation of equity-based compensation however presents a different kind of agency problem between shareholders and managers regarding the optimal level of firm risk-taking. Equity-based compensation gives managers incentives to work harder but on the other hand, this compensation structure may result in an undesirable outcome which is that most managers, whose wealth now is concentrated mostly in the company stock, will less likely to invest in risky projects in order to maintain the value of his undiversified wealth. Many studies on corporate risk-taking document how executive compensation affects corporate risk-taking and find that equity-based compensation encourages CEOs to take more risks. Coles, Daniel, and Naveen (2006) show that CEOs whose compensation is more sensitive to the firm's volatility (CEOs with higher vega) are more likely to make more risky investment decisions.

Another possible link connecting CEO and firm risk-taking is through CEO's personal risk-taking preference. This growing branch of literature includes the study of Bertrand and Schoar (2003) who establish the relation between manager's personal characteristics and corporate behavior and performance. Malmendier and Tate (2005) establish a relation between managerial overconfidence and corporate investment decisions. They find evidence that overconfident CEOs tend to overinvest. Frank and Goyal (2009) find that CEO fixed effect matters in the choice of corporate leverage, both in the level of leverage and also in the leverage adjustment speed. They also argue that the CEO fixed effect is a first order effect in corporate leverage. Using a sample of manufacturing and technology firms, Galasso and Simcoe (2011) find that overconfident CEOs are more willing to invest in more risky projects and to lead their firms to a change in direction. An interesting paper by Cain and McKeon (2011) finds that CEOs with high personal risk-taking (represented by CEOs who possess small airplane pilot license) increase firm's overall risk and those CEOs are associated with value increasing acquisitions. Cronqvist, Makhija, and Yonker (2012) find that CEO's personal leverage is positively correlated with their firms' choice of capital structure. In other words, they find that CEOs with more conservative personal leverage (those with lower mortgage) are more likely to implement more conservative capital structure for their firms.

Firms hire their CEOs either from outside of the firm (outsider CEO) or promote a senior executive from within the firm (insider CEO). Murphy and Zabojnik (2007) document that although CEO turnovers are still dominated by insider CEOs, firms are more likely to fill their CEO positions through external hires than through internal promotions. The board of directors is more likely to hire outside CEOs when the firm struggles and needs a change in direction. Ferris, Wei, and Zhang (2007) find that in general, firms that perform poorly hire a new CEO from outside of the company while firms that perform better hire a new CEO from inside of the company. They also find that outsider CEOs are paid more than insider CEOs and mainly through incentive pay, which is consistent with findings of Murphy (2002).

Unfortunately, there is not much explanation why firms are more likely to hire outsider CEOs now. It is a puzzling phenomenon because outsider CEOs command higher compensation than insider CEOs and hiring outsider CEOs could discourage senior executives in the firm to do well. The CEO position is usually viewed as the ultimate reward for a senior executive who does well. If the board removes the possibility that a well performing executive could become a CEO, the board does not provide the firm senior executives with the incentives to do well. Thus, if the board still hires outsider CEOs (a practice that requires higher financial investments and consequently presents disincentive for senior executives to do well), then there must be some benefits that outweigh the aforementioned costs.

It is reasonable to think that firms hire outsider CEOs because they produce better performance than insider CEOs. However, this argument is not in consistent with findings of Zhang and Rajagopalan (2010). They find that outsider CEOs outperform insider CEOs in the early part of their tenure and significantly underperform insider CEOs in the long run. Therefore, considering the higher costs to hire outsider CEOs, it is perplexing why firms hire outsider CEOs.

In this paper, we examine how a CEO career origin affects firm risk-taking and whether the reason the board of directors hires their new CEOs from outside of the firm is because of the new CEO's personal risk-taking preference.

We find that outsider CEOs make more risky investment decisions (invest more in R&D, invest less in capital expenditures, and use more leverage) than insider CEOs. One can argue that outsider CEOs take more risks because of their compensation structure has more equity-based portion and that kind of compensation structure encourages CEOs to take more risk. However, after controlling for compensation sensitivity variables such as delta and vega, we find that outsider CEOs are still significantly more risk-taking than insider CEOs. Our results suggest that outsider CEOs are more risk-taking than insider CEOs and the board of directors, despite the higher costs of hiring outsider CEOs, may very well hire outsider CEOs because of the risk-taking reason. Our results also suggest that the CEO compensation structure difference (between outsider and insider CEOs) is not the reason why outsider CEOs take more risk than insider CEOs.

In addition to examining the relation between CEO career origin and the *level* of the firm risk-taking, we also examine the relation between CEO career origin and the *change* in the firm risk-taking years after the CEO turnover. We find that on average, outsider CEOs increase the normalized R&D ratio by 1.77% in the four year period (from one year before the CEO turnover to three years after the turnover), while the insider CEOs increase the normalized R&D ratio only by 0.63% during the same time period. This also helps explain why the board is more likely to hire outsider CEOs when the firms struggle.

Our paper contributes to the CEO turnover and corporate risk-taking literature in three different ways. First, we show that a CEO career origin matters in corporate risk-taking and we extend the literature on the relation between CEO personal characteristic and corporate risk-taking. Second, we provide an alternative explanation as to why firms hire more outsider CEOs now than before. We argue that outsider CEOs do not make more risky investment decisions because of their compensation structure but they make more risky investment decisions because they are more risk-taking than insider CEOs and that may be the reason firms hire more outsider CEOs now than before. Third, we find mixed results about the relation between a CEO career origin and corporate risk-taking. Interestingly outsider CEOs increase the risk of the firm by investing more in R&D and less in capital expenditure, but *not through* increased leverage.

The rest of the paper continues as follows. In section 2, we describe our data collection and how we construct our main variables, including the CEO and firm characteristic variables. Section 3 presents all the results and section 4 concludes.

Data and Sample Description

We construct our sample using CEO data from the Standard & Poor's Execucomp database between 1992 and 2010. We identify a CEO if an executive has a title of "CEO" (i.e., if the Execucomp's "titleann" equals "CEO" or "Chief Executive Officer" or "Principal Executive Officer") and if the executive is clearly identified as a CEO by Execucomp (when Execucomp's CEO identifier "ceoann" equals "CEO"). Following Coles, Daniel, and Naveen (2006), we also include executives who are not identified by Execucomp as CEOs but they actually assume the CEO position. We exclude utilities firms (SIC 4900-4999) and financial services firms (SIC 6000-6999) from our sample. We obtain 24,649 CEO-year observations during this time period.

As we are interested in finding how a CEO career origin affects firm risk-taking, we furthermore search for CEO turnovers that took place during this time period and classify the new CEOs into 2 categories: insider CEOs and outsider

CEOs. We define insider CEO as a CEO who is promoted from within the company or if the new CEO had worked for the company at least 1 year before he was appointed as the new CEO. We define outsider CEO as a new CEO who is hired from outside the company or if the new CEO had only worked for the company for less than 1 year before he was appointed to be the new CEO. We then construct a dummy variable, *OUTSIDER*, which equals 1 for outsiders and 0 for insiders. After excluding all interim CEOs, we find 3,006 CEO turnovers within this time period (among those turnovers, there are 2,212 insider CEOs and 794 outsider CEOs).

We collect CEOs' personal data such as age and gender, the CEO compensation data such as salary, bonus, and equity compensation. We create the CEO/founder dummy, which takes a value of 1 if a CEO is also the founder of the company and 0 otherwise. We find 203 CEOs who are also the founders of their firms. We create the CEO/chairman dummy, which takes a value of 1 if a CEO is also the chairman of the board and 0 otherwise. We find 1,604 CEOs who are also the chairman of the board.

For the accounting data, we collect data from Compustat database over the same time period (1992-2010). The accounting data are deflated to constant 2000 dollars using the GDP deflator and winsorized at the 0.50% level at both tails to minimize the effect of outliers and mis-recorded data. Folowing Coles, Daniel, and Naveen (2006), we use three risk-taking variables: (1) *R&D*, defined as the ratio of research and expenditures expenses divided by assets (2) *CAPEX*, defined as the ratio of net capital expenditures (capital expenditures minus sales of property, plant, and equipment) divided by assets (3) *LEVERAGE*, defined as the ratio of long term debt plus debt in current liabilities divided by assets. For independent variables, we use the *OUTSIDER* dummy variable and control variables (firm characteristic variables and compensation sensitivity variables like delta and vega). For personal CEO characteristics, we collect age, tenure, sex, and two dummy variables: the chairman and founder dummies. The chairman dummy takes a value of 1 if the CEO is also the chairman of the board in that year and 0 otherwise. The founder takes a value of 1 if the CEO is the founder of the company. For compensation variables, we calculate delta and vega. Delta is the change in the dollar value of the CEO's wealth for one percentage point change in the company's stock price while vega is defined as the change in the dollar value of the CEO's wealth for a 0.01 change in the annualized standard deviation of stock returns. To calculate delta and vega, we follow Guay (1999) and Core and Guay (2002).

Results

We test the mean difference of risk-taking measures (such as R&D, capital expenditure, book value of leverage, and market value of leverage) implemented by outsider and insider CEOs ten years after the CEO turnover. We examine whether the difference between average risk-taking implemented by outsider and insider CEOs is statistically different from zero. From table 3, we find results that are consistent with our expectation that the average of risk-taking measures implemented by outsider CEOs is statistically different from that implemented by insider CEOs and the difference is statistically significant at 1% level in every year for the ten years after the CEO turnover. We also examine the hypothesis that outsider CEOs are more of a risk-taker than insider CEOs and again we also find convincing evidence that outsider CEOs implement more risky investment decisions than insider CEOs (indicated by higher R&D and lower capital expenditures) and that the results are statistically significant at 1% level. Interestingly, we find that the average leverage (measured either by book value of leverage or market value of leverage) is lower for outsider CEOs than insider CEOs. We discuss this specific finding about leverage in the univariate analysis section.

We find that the CEO origin dummy is positively correlated with R&D and negatively correlated with capital expenditures. These results both imply that outsider CEOs take more risk than insider CEOs. Interestingly, we find that there is a negative correlation between the CEO origin dummy and measure of leverage (measured either by book value of leverage or by market value of leverage) and this negative correlation is significant at 5% level. The negative sign of the correlation is the opposite of what we expect from the correlation between CEO origin dummy and leverage and this negative correlation is consistent with the finding that the average leverage ratio implemented by outsider CEOs is lower than that of the insider CEOs in the previous section. As unexpected as the negative correlation is, this particular finding is in line with the finding of Frank and Goyal (2009) who argue that in the event that there is a forced CEO turnover, the company leverage will be elevated one year before the turnover and outsider new CEO will reduce the elevated level of leverage by more than insider new CEOs will. Combining the finding of Frank and Goyal (2009) with the finding that financially troubled firms are more likely to hire outsider CEOs to change the direction of the firm, we find an interesting result that although outsider CEOs take more risks than insider CEOs, they increase the risk of the company through higher R&D expenditure and lower capital expenditure investment and *not through* higher leverage. This finding contributes to the literature of corporate risk-taking that although outsider CEOs are more of a risk taker, they increase risk through investing more in R&D and less in capital expenditures but *not* through higher leverage.

To control for other CEO personal and compensation variables as well as for firm characteristic variables, we examine the role of a CEO career origin on corporate risk-taking utilizing the multivariate analysis. We run multivariate regression analysis to examine whether and how a CEO career origin affects the corporate risk-taking while controlling

for CEO compensation sensitivity variables (such as delta and vega) and for firm characteristic variables. The regressions are as follows:

$$R\&D_{i,t} = \alpha + \lambda(CEO\ career\ origin\ dummy)_{i,t-1} + \beta(CEO\ personal\ characteristc)_{i,t-1} + \gamma(CEO\ pay\ characteristic)_{i,t-1} + \theta(Firm\ Characteristic)_{i,t-1} + \varepsilon_{i,t}$$

$$CAPEX_{i,t} = \alpha + \lambda(CEO\ career\ origin\ dummy)_{i,t-1} + \beta(CEO\ personal\ characteristc)_{i,t-1} + \gamma(CEO\ pay\ characteristic)_{i,t-1} + \theta(Firm\ Characteristic)_{i,t-1} + \varepsilon_{i,t}$$

$$BkLev_{i,t} = \alpha + \lambda(CEO\ career\ origin\ dummy)_{i,t-1} + \beta(CEO\ personal\ characteristc)_{i,t-1} + \gamma(CEO\ pay\ characteristic)_{i,t-1} + \theta(Firm\ Characteristic)_{i,t-1} + \varepsilon_{i,t}$$

$$MktLev_{i,t} = \alpha + \lambda(CEO\ career\ origin\ dummy)_{i,t-1} + \beta(CEO\ personal\ characteristc)_{i,t-1} + \gamma(CEO\ pay\ characteristic)_{i,t-1} + \theta(Firm\ Characteristic)_{i,t-1} + \varepsilon_{i,t}$$

In table 2 column (1), we start with the pooled OLS regression using firm fixed effects. We find that all coefficients are consistent with the corporate risk-taking literature. Then, in column (2) we add the CEO career origin dummy into the regression. We find that the CEO origin effect is statistically significant at 1% level and this result implies that outsider CEOs make more R&D investment than insider CEOs. In column (3), we rerun the regression using the clustered standard errors by year and by industry and we still find similar results.

In table 3, we examine how the CEO career origin affects the firm's capital expenditure. We find that outsider CEOs invest less in capital expenditure than insider CEOs. This finding is consistent even after we cluster the standard errors by year and by industry. We find that outsider CEOs use more leverage than insider CEOs. This finding is not really inconsistent with the findings we find from the mean difference test and the univariate analysis. It simply means that outsider CEOs use less leverage in the first several years in their tenure and increase the leverage over time.

Conclusions

In this paper, we examine how a CEO career origin affects firm risk-taking and whether the reason the board of directors hires their new CEOs from outside of the firm is because of the new CEO's personal risk-taking preference. We find that outsider CEOs make more risky investment decisions (invest more in R&D, invest less in capital expenditures, and use more leverage) than insider CEOs. One can argue that outsider CEOs take more risks because of their compensation structure has more equity-based portion and that kind of compensation structure encourages CEOs to take more risk. However, after controlling for compensation sensitivity variables such as delta and vega, we find that outsider CEOs are still significantly more risk-taking than insider CEOs. Our results suggest that outsider CEOs are more risk-taking than insider CEOs and the board of directors, despite the higher costs of hiring outsider CEOs, may very well hire outsider CEOs because of the risk-taking reason. Our results also suggest that the CEO compensation structure difference (between outsider and insider CEOs) is not the reason why outsider CEOs take more risk than insider CEOs.

In addition to examining the relation between CEO career origin and the *level* of the firm risk-taking, we also examine the relation between CEO career origin and the *change* in the firm risk-taking years after the CEO turnover. We find that on average, outsider CEOs increase the normalized R&D ratio by 1.77% in the four year period (from one year before the CEO turnover to three years after the turnover), while the insider CEOs increase the normalized R&D ratio only by 0.63% during the same time period. This also helps explain why the board is more likely to hire outsider CEOs when the firms struggle.

Our paper contributes to the CEO turnover and corporate risk-taking literature in three different ways. First, we show that a CEO career origin matters in corporate risk-taking and we extend the literature on the relation between CEO personal characteristic and corporate risk-taking. Second, we provide an alternative explanation as to why firms hire more outsider CEOs now than before. We argue that outsider CEOs do not make more risky investment decisions because of their compensation structure but they make more risky investment decisions because they are more risk-taking than insider CEOs and that may be the reason firms hire more outsider CEOs now than before. Third, we find mixed results about the relation between a CEO career origin and corporate risk-taking. Interestingly outsider CEOs increase the risk of the firm by investing more in R&D and less in capital expenditure, but *not through* increased leverage.

References

Bebchuk, Lucian and Yaniv Grinstein, 2005. The growth of executive pay. *Oxford Review of Economic Policy* 21, 283-303.

Bertrand, Marianne and Antoinette Schoar, 2003. Managing with style: the effect of managers on firm policies. *Quarterly Journal of Economics* 118, 1169-1208.

Cain, Matthew and Stephen McKeon, 2011. Cleared for takeoff? CEO personal risk-taking and corporate policies. Working paper.

Coles, Jeffrey, Naveen Daniel, and Lalitha Naveen, 2006. Managerial incentives and risk-taking. *Journal of Financial Economics* 79, 431-468.

Core, John and Wayne Guay, 2002. Estimating the value of employee stock option portfolio and their sensitivities to price and volatility. *Journal of Accounting Research* 40, 613-630.

Cronqvist,Henrik, Anil Makhija, and Scott Yonker, 2012. Behavioral consistency in corporate finance: CEO personal and corporate leverage. *Journal of Financial Economics* 103, 20-40.

Datta, Sudip, Mai Iskandar-Datta, and Kartik Raman, 2001. Executive compensation and corporate acquisition decisions. *Journal of Finance* 56, 2299-2336.

Eisenmann, Thomas, 2002. The effects of CEO equity ownership and firm diversification on risk taking. *Strategic Management Journal* 23, 513-534.

Ferris, Stephen, Zuobao Wei, and Shaorong Zhang, 2007. CEO succession, incentives and firm performance: insiders versus outsiders. Working paper.

Frank, Murray and Vidhan Goyal, 2009. Corporate leverage: how much do managers really matter? Working paper.

Galasso, Alberto and Timothy Simcoe, 2011. "CEO overconfidence and innovation" Working paper.

Guay, Wayne, 1999. The sensitivity of CEO wealth to equity risk: an analysis of the magnitude and determinants. *Journal of Financial Economics* 53, 43-71.

Huson, Mark, Paul Malatesta, and Robert Parrino, 2004. Managerial succession and firm performance. *Journal of Financial Economics* 74, 237-275.

Low, Angie, 2009. Managerial risk-taking behavior and equity-based compensation. *Journal of Financial Economics* 92, 470-490.

Malmendier, Ulrike and Geoffrey Tate, 2005. CEO overconfidence and corporate investments. *Journal of Finance* 60, 2661-2700.

Murphy, Kevin, 2002. Explaining executive compensation: managerial power vs. perceived cost of stock options. *University of Chicago Law Review* 69, 847-863.

Murphy, Kevin and Jan Zabojnik, 2007. Managerial capital and the market for CEOs. Working paper.

Sanders, Gerard, 2001. Behavioral responses of CEOs to stock ownership and stock option pay. *Academy of Management Journal* 44, 477-492.

Wang, Lingling, 2009. CEO employment history and risk-taking in firm policies. Working paper.

Zhang, Yan and Nandini Rajagopalan, 2010. Once an outsider, always an outsider? CEO origin, strategic change, and firm performance. *Strategic Management Journal* 31, 334-346.

Table 1: R&D regression

Variable	Pooled OLS (1)	Pooled OLS (2)	Clustered SE by Year (3)
CEO Origin		0.005***	0.005***
		(4.26)	(3.71)
CEO Characteristic			
Age	0.0002**	0.0001*	0.0001
	(2.07)	(1.82)	(1.29)
Tenure	-0.0004***	-0.0004***	-0.0004***
	-(5.05)	-(4.98)	-(4.45)
Sex	0.0134***	0.0133***	0.01326883***
	(3.68)	(3.63)	(4.58)
CEO Pay Characteristic			
Delta	-0.0012***	-0.0011***	-0.0011***
	-(2.55)	-(2.45)	-(2.9)
Vega	0.0044***	0.0045***	0.0045***
	(3.78)	(3.89)	(2.38)
Cash Compensation	0.0027***	0.0027***	0.0027***
	(4.25)	(4.14)	(2.62)
Firm Characteristic			
Market to Book	0.0112***	0.0112***	0.0112***
	(27.55)	(27.59)	(6.68)
Sales Growth	0.0018***	0.0018***	0.0018***
	(3.53)	(3.37)	(2.16)
Log(Sales)	-0.0051***	-0.0049***	-0.0049***
	-(12.63)	-(12.29)	-(8.47)
ROA	-0.2199***	-0.2193***	-0.2193***
	-(42.26)	-(42.16)	-(10.76)
Surplus Cash	0.1485***	0.1468***	0.1468***
	(41.45)	(40.75)	(12.81)
Constant	0.0204***	0.02***	0.02***
	(3.15)	(3.1)	(2.31)
2-digit SIC controls	Yes	Yes	Yes
Observations	14703	14703	14703
Adjusted R^2	0.3808	0.3815	0.382

R&D is the research and development expenses (XRD) scaled by AT. Missing R&D has been set to zero. CEO Career Origin Dummy is an indicator variable that takes the value of 1 for outsider CEO and 0 for insider CEO. CEO personal characteristic, CEO pay characteristic, and Firm characteristic are explained in Appendix A. * is 10% significance level, ** is 5% significance level, and *** is 1% significance level.

Table 2: CAPEX regression

Variable	Pooled OLS (1)	Pooled OLS (2)	Clustered SE by Year (3)
CEO Origin		-0.0061***	-0.0061***
		-(5.86)	-(6.41)
CEO Characteristic			
Age	-0.0002***	-0.0002***	-0.0002***
	-(3.11)	-(2.77)	-(3.29)
Tenure	0.0003	0.0002***	0.0002***
	(3.68)	(3.59)	(4.3)
Sex	0.0094***	0.0096***	0.0096***
	(2.94)	(3.01)	(2.89)
CEO Pay Characteristic			
Delta	0.0021***	0.0021***	0.0021***
	(5.24)	(5.1)	(3.46)
Vega	-0.0063	-0.0065	-0.0065
	-(6.18)	-(6.34)	-(5.01)
Cash Compensation	-0.0009**	-0.0009*	-0.0009*
	-(1.67)	-(1.51)	-(1.06)
Firm Characteristic			
Market to Book	0.0013***	0.0013***	0.0013**
	(3.75)	(3.72)	(1.64)
Sales Growth	-0.0014***	-0.0013***	-0.0013
	-(3.15)	-(2.92)	-(1.4)
Log(Sales)	-0.0046***	-0.0047***	-0.0047***
	-(13.19)	-(13.59)	-(7.88)
ROA	0.1285***	0.1278***	0.1278***
	(28.34)	(28.21)	(9.63)
Surplus Cash	-0.0478***	-0.0457***	-0.0457***
	-(15.3)	-(14.57)	-(7.94)
Constant	0.0815***	0.0819***	0.0819***
	(14.44)	(14.53)	(12.72)
2-digit SIC controls	Yes	Yes	Yes
Observations	14703	14703	14703
Adjusted R^2	0.0975	0.0995	0.1003

CAPEX is the capital expenditure (CAPX) scaled by AT. CEO Career Origin Dummy is an indicator variable that takes the value of 1 for outsider CEO and 0 for insider CEO. CEO personal characteristic, CEO pay characteristic, and Firm characteristic are explained in Appendix A. * is 10% significance level, ** is 5% significance level, and *** is 1% significance level.

UNCW

UNIVERSITY of NORTH CAROLINA WILMINGTON

DEPARTMENT of ECONOMICS AND FINANCE

601 South College Road
Wilmington, NC 28403-5945