IN THIS ISSUE

Macroeconomic Determinants of Economic Growth: A Cross-Cohort Study of Anglophone, Francophone and Hispanophone Countries...............................Abdullah M. Khan

Presidential Elections and Industry Stock Returns: A Test of Market Efficiency ..........................................................Patrick Gobran and Frank Bacon

Internet Trolling in Social Networking Sites: A Preliminary Investigation of Undergraduate Student Victimization....................................................Carl Case and Darwin King

A Path from Job Autonomy to Organizational Citizenship Behavior: The Role of Perceived Organizational Politics as Mediator ....................................Tae Seok Yang, Alankrita Pande, Yin-Chi Liao and Joseph J. Dobson

Effects of Employee Heterogeneity on Firm Performance: Measuring The Impact of International Diversity in the National Basketball Association ...............Mayo-Smith, Fenn and Parco

Accounting History in Perspective: Uniform CPA Exam Turns 100 ..........................................................King, Case and Senecker

An Investigation of Investor Reaction to Corporate Social Responsibility Motivation and Ethical Position in an Environmental Context ........................Karen .  Green and Melloney Simerly

Seasonality in the Monthly Returns of Large Stocks: 1926 To 2013 ..........................................................Shaikh A. Hamid

When Rivals Purchase: This Means War! Or Does it? .................................................................Burns, Hutchins and Mathisen

Does Unwinding Carry Trade Lead to Profitable Reverse Carry Trade? The Case of Yen .................................................................Minje Jung

The Silver Tsunami: Evaluating the Impact of Population Aging in the U.S. ..........................................................Henderson, Maniam and Leavell

Incentivizing Management Discretionary Philanthropy: Social Profit Credits ..........................................................Minutolo, Mills and Stakeley

Roth IRAs and Qualified Charitable Distributions .............................................................Sheldon R. Smith and Lynn R. Smith
The Journal of Business and Behavioral Sciences is a publication of the American Society of Business and Behavioral Sciences (ASBBS). Papers published in the Journal went through a blind review process prior to acceptance for publication. The editors wish to thank anonymous referees for their contributions.

The national annual meeting of ASBBS is held in Las Vegas in March of each year and the international meeting is held in May/June of each year. Visit www.asbbs.org for information regarding ASBBS.
<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Macroeconomic Determinants of Economic Growth: A Cross-Cohort Study of Anglophone, Francophone and Hispanophone Countries</td>
<td>Abdullah M. Khan</td>
<td>4</td>
</tr>
<tr>
<td>Presidential Elections and Industry Stock Returns: A Test of Market Efficiency</td>
<td>Patrick Gobran and Frank Bacon</td>
<td>21</td>
</tr>
<tr>
<td>Internet Trolling in Social Networking Sites: A Preliminary Investigation of Undergraduate Student Victimization</td>
<td>Carl Case and Darwin King</td>
<td>32</td>
</tr>
<tr>
<td>A Path from Job Autonomy to Organizational Citizenship Behavior: The Role of Perceived Organizational Politics as Mediator</td>
<td>Yang, Pande, Liao and Dobson</td>
<td>44</td>
</tr>
<tr>
<td>Effects of Employee Heterogeneity on Firm Performance: Measuring The Impact of International Diversity in the National Basketball Association</td>
<td>Mayo-Smith, Fenn and Parco</td>
<td>57</td>
</tr>
<tr>
<td>Accounting History in Perspective: Uniform CPA Exam Turns 100</td>
<td>King, Case and Senecker</td>
<td>70</td>
</tr>
<tr>
<td>An Investigation of Investor Reaction to Corporate Social Responsibility Motivation and Ethical Position in an Environmental Context</td>
<td>Karen Y. Green and Melloney C. Simerly</td>
<td>85</td>
</tr>
<tr>
<td>Seasonality in the Monthly Returns of Large Stocks: 1926 To 2013</td>
<td>Shaikh A. Hamid</td>
<td>103</td>
</tr>
<tr>
<td>When Rivals Purchase: This Means War! Or Does it?</td>
<td>Burns, Hutchins and Mathisen</td>
<td>122</td>
</tr>
<tr>
<td>Does Unwinding Carry Trade Lead to Profitable Reverse Carry Trade? The Case of Yen</td>
<td>Minje Jung</td>
<td>140</td>
</tr>
<tr>
<td>Incentivizing Management Discretionary Philanthropy: Social Profit Credits</td>
<td>Minutolo, Mills and Stakeley</td>
<td>170</td>
</tr>
<tr>
<td>Roth IRAs and Qualified Charitable Distributions</td>
<td>Sheldon Smith and Lynn Smith</td>
<td>182</td>
</tr>
</tbody>
</table>
MACROECONOMIC DETERMINANTS OF ECONOMIC GROWTH: A CROSS-COHORT STUDY OF ANGLOPHONE, FRANCOPHONE AND HISPANOPHONE COUNTRIES

Abdullah M. Khan
Claflin University

ABSTRACT

The focus of this paper is to analyze the influence of several macroeconomic determinants on per capita GDP growth of countries from three language cohorts (Anglophone, Francophone, and Hispanophone). Differences in per capita real GDP across countries are stylized facts in economic growth literature. Even though the empirical literature varies widely in attributing such differentials to any specific set of factors, a strand of growth literature finds institutional differences in colonial legacies as a major contributing factor to growth differences across countries. In this study, we test the hypothesis that there are statistically significant differences in per capita real GDP across countries with three dominant languages. We construct three language cohorts of countries to serve as proxies for institutional differences from colonial legacies, as well as, lingually driven differences in economic behaviors of agents. The GLS panel regression results confirm that Anglophone countries’ per capita GDP is higher than that of Hispanophone and Francophone countries.

Keywords: Economic growth, Anglophone, Francophone, Hispanophone

INTRODUCTION

A growing strand of empirical studies find institutional legacies, in addition to other socio-economic variables, influencing factors of economic growth outcomes across countries and regions (e.g., Acemoglu, Johnson, and Robinson, 2006, 2005, 2002, 2001; Acemoglu, Johnson, Robinson, and Thaicharoen, 2003; Ali, 2003; Bertocchi and Canova, 2002; Landes, 1998; Mahoney, 2001; Grier, 1999; Rodrick, Subramanian, and Trebbi, 2004). The colonial institutional legacy varied in terms of location of the settlements, property rights and contractual rights determination and enforcement, recruitment of settlers and workers, prevalence of indentured servitudes, extent of autonomy given to specific colonial rulers from the mainland, governance structure, religious practices, language spoken by inhabitants to name a few (Acemoglu and Johnson, 2005; Besley, 1995; Chen, 2013; Hall and Jones, 1999; Johnson and Woodruff, 2002; Knack and Keefer, 1995; La Porta, Lopez-de-Silanes, Shleifer, and Vishny, 1999; North, Summerhill, and Weingast, 2000; La Porta, Lopez-de-Silanes, Shleifer, and Vishny, 1998;
Western European colonial rulers were the main imperial forces with that once dominated the global landscape from beginning of the fourteenth century up to middle of twentieth century. The Portuguese were one of the earliest founders of colonial empires with trading posts around the world including parts of India, Brazil, few Middle Eastern ports, Timor etc. The Portuguese empire had its peak in the fifteenth century and declined gradually thereafter. The Spanish, Dutch, English and French rulers controlled vast colonial territories in the Asia, Africa, North America, South America and the Caribbean islands. The ubiquity of colonial rules can be put on perspectives by recalling the historical fact that 125 countries out of current 192 countries that are members of the United Nations’ were once subjugated as colonies of few European nations (Abernethy, 2000) and by 1930s, colonies and former colonies represented 84.6 percent of land surface of the world (Fieldhouse, 1967). The British Empire emerged as the largest empire through the eighteenth and nineteenth century in terms of hegemonic influence and land coverage of colonies and population thereof. The imperial powers were keen on introducing their native languages in conducting official businesses in colonial settlements essentially making languages part of colonial governance strategies (see Bokamba, 1991, Hamel, 2005; Phillipson, 1992, Phillipson, 1997). For most instances, the residents of former colonies continue to use, either officially or otherwise, the languages once were introduced by their colonial rulers. Using a game theoretic approach Arias and Girod (2011) studied 444 sub-national territories in the Americas for both the pre-colonial and post-colonial periods across different European colonial regimes and reported preexisting differences in economic and political development that predate era of colonization. This finding lends credence to the possibility that pre-colonial institutional differences might have been augmented due to the colonial institutional practices.

In this paper, we attempt to extend existing literature the area of cross-country comparison of economic growth mainly by using three language cohorts of countries as proxies of institutional differences and employ a panel regression model to test if there are statistically significant differences across them. In the light of several recent empirical studies, we contend that in addition to serving as proxies for institutional differentials, language cohorts also serve as embodiment of differentials in human economic behaviors shaped party by lingual structures or expressions regarding temporal emphasis on present vs. future (e.g., Chen, 2013; Cronqvist and Siegel, 2015; Dahl and Velupillai, 2011; Dahl 2000). The remaining sections of the paper is organized as follows: section 2 presents a review of contemporary literature on economic growth, section 3 contains a discussion on data, variable, and regression model, section 4 summarizes results, and section 5 concludes.
Acemoglu, Johnson, and Robinson (2001) reports evidence that economic and political institutions account for about seventy five percent of the differences in per capita income across countries that became independent from various colonial empires. Lee and Schultz (2012) comparatively study economic development of Cameroon and report evidence of faster economic development in the British colonial part of Cameroon than the French colonial part. Brunt (2011) reports a natural experiment from South Africa involving agricultural produce data from British and Dutch colonial era and found evidence that property rights enhances productivity and therefore economic growth. Several other researchers (e.g., Alexiou, Tsaliki, and Osman, 2014; Pääkkönen, 2010; Paul, Torres, and Felihooven, 2009; Osman, Alexiou, and Tsaliki, 2012) reported impact of institutions on economic growth-related outcomes.

Acemoglu (2009) provides a comprehensive review of contemporary empirical work on aspects of economic growth including cross country convergence and divergence of income in the recent decades. Acemoglu, Johnson, and Robinson (2005) reports positive impact of institutions as a significant catalyst of economic growth across various countries including the ones that used to share similar socio-cultural heritages but are under different institutional arrangements at present (e.g., North Korea and South Korea). Rodrik, Subramanian, and Trebbi (2004) find strong evidence that institutions matter for economic growth. Tadei (2014) found evidence that regional development was adversely correlated for regions of French colonial Africa with agricultural product prices lower than world market prices.

Earlier empirical work on growth convergence include Barro (1991), and Barro and Sala-i-Martin (1992a, 1992b), Mankiw, Romer and Weil (1992), Islam (1995) among others. Mankiw, Romer, and Weil (1992) studied cross-country economic growth data and found evidence in support of growth rate convergence among developed vs. developing countries at a rate predicted by augmented Solow model (Solow, 1956). Barro (1991) reported positive influence of human capital on economic growth, and that, poorer countries with low capital-to-labor ratio grew faster than their richer counterparts during the study period. Barro (1997) is a study of convergence conditions with findings that that poorer countries may grow faster than their richer counterpart if the poorer countries’ per capita GDP is further below their steady state-level of per capita GDP relative to richer countries’ per capita GDP distance from their own steady state.

An emerging strand of interdisciplinary literature looking at relationship between spoken and written languages and economic behaviors of their users. Several studies (Dahl and Velupillai, 2011; Dahl, 2000; Dahl, 1985; Dahl and Dienes, 1984) report distinctive structures of temporal expressions across languages which, according to some other studies, (e.g., Chen, 2013; Sutter,
Angerer, Glätzle-Rützler, and Lergetporer, 2015) have contributed to differential economic behaviors of users these languages. The World Atlas of Language Online is a comprehensive database about languages around the world and their distinctive features. Chapter sixty-seven of this online database distinguishes world languages in terms of whether a given language culture warrants inflectional marking of future / non-future distinction or not. Also, these studies contend that languages can be found in a continuum in terms of their practices of using explicit inflectional marking of future/non-future distinction. According to Dahl and Velupillai’s (2011) chapter 67 of the World Atlas of Structures Online (http://wals.info/) English language has ‘no inflection future’ (source: http://wals.info/languoid/lect/wals_code_eng), in French language ‘inflectional future exists’ (http://wals.info/languoid/lect/wals_code_fre), and in Spanish Language ‘inflectional future exists’ (http://wals.info/languoid/lect/wals_code_spa). On the World Atlas of Language Structures Online website after accessing each of these three weblinks one should click on “Fid” no. 67A to read whether inflectional future marker exists or not for the language selected For English language, this can also be read in Dahl (1985, p. 122), for French language’s inflectional future marker please see Dahl (1985, p. 170), and for Spanish language please refer to Dahl (1985, p. 171). Chen (2013) adopts Dahl and Velupillai’s (2011) framework of countries with different inclinations for inflectional marking of future / non-future distinctions and empirically test to find evidence that inflectional markings differences across language groups may impact economic behavior of the users of these languages. In this paper, we test the hypothesis that if there exists statistically significant difference in economic growth outcomes across countries grouped into three separate languages cohorts. We contend that the differences in growth comes from institutional legacies and from differences in economic behavior of residents induced by differences in inflectional marking of future time reference. We contend that the more a language is prone to or explicit about inflectional marking of future vs. non-future (relative to another language that does not require use of distinct future tense indicating verbs) the less pragmatic the users of that language will be about future and vice versa. According to the World Language Structure database chapter 67 (Dahl, 1985; Dahl and Velupillai, 2011) English is not particularly known as inflection marking of future / non-future, but French and Spanish languages both are particularly prone to inflection marking of future /non-future.

The difference in lingual expressions in this context is implying that statements in a language regarding future events or actions (anticipated, scheduled or unscheduled) may contain ‘more explicit’ futuristic verbs than in another language. The World Atlas of Languages Structures Online has conducted extensive analysis and created language structure maps that explains for which languages inflectional future tense exist and for which languages it does not exist. For French language is more explicit in terms of inflection of future than Finnish language. In Finnish language ‘It is cold today’ (word by word translation of
equivalent Finnish sentence) and ‘it is cold tomorrow’ both are grammatically correct and commonly acceptable. But, in French language ‘It will be cold tomorrow’ is the correct statement and ‘it is cold tomorrow’ will be considered outright wrong sentence to a French speaker. According to Dahl (1985) Finnish is a language with no inflection future marking because Finnish language does not explicitly separate future from present by using future tense indicating verbs. On the other hand, for the reasons given above, Dahl (1985) categorizes French as a language with inflection future marker because explicitly separate future from present by using future tense indicating verbs. There studies that contend that such differences in expressions regarding present vs. future can impact economic behaviors of the speakers.

**DATA, METHOD, AND VARIABLE CONSTRUCTION**

The data was collected for the years 1990 through 2013 from various sources including World Bank Database, Federal Reserve Bank of St. Louis database, CIA’s world Fact book, and IMF’s data base. The macroeconomic variables that We use in this study are per capita GDP (in purchasing power parity dollars), trade openness, internet usage, ethnic diversity, electricity consumption, tertiary enrolment (a proxy for human capital), and foreign direct investment inflow as percent of GDP. We also use dummy variables for three languages -cohorts: Anglophone, Francophone, and Hispanophone countries.

The baseline generalized least square (GLS) panel regression model used in this study is as follows:

\[
GDP\ growth\ rate\ \ ct = B_1 + B_1 \times (ln\ R&D\ share\ of\ GDP)_{ct} + B_2 \times (ln\ Trade\ Openness)_{ct} + B_3 \times (ln\ Internet\ Usage)_{ct} + B_4 \times (ln\ Ethnic\ Diversity)_{ct} + B_5 \times (ln\ Electricity\ Consumption)_{ct} + B_6 \times (ln\ Tertiary\ Enrolment)_{ct} + B_7 \times (ln\ Foreign\ Investment\ Inflow\ as\ percent\ of\ GDP)_{ct} + Error\ Terms.
\]

In the variant regression model, we add two dummy variables for two language cohorts.

In the above specification, the subscripts “c” indicates country and subscript “t” indicates year for the panel dataset where t=1997, 1998,…,2013. The time coverage for the paper (1997-2013) was data availability driven. On the World bank database GDP data before 1997 is more marred with missing values we decided to include data from 1997 to 2013 as in this time frame most of the countries have required data available. In the regression model presented above, variable ‘trade openness’ was calculated as export plus import as percent of GDP. ‘Internet usage’ variable implies percent of people with access to internet. The Ethnic diversity variable is created as a ratio of percentage sum of all minority groups to percentage of largest ethnic group. For example, if share of Niger’s largest ethnic group (Haoussa) is 55.4 percent of the population of Nigeria, then the percent sum of minority ethnic groups’ combined population will be 44.6 percent (i.e., 100 percent minus 55.4 percent), and the diversity index for Niger
will be 0.81 (i.e., 44.6 / 55.4). Similarly, if we are to calculate Sri Lanka’s diversity index, it would be 0.34 and Niger will be considered as more ethnically diverse (i.e., 0.81 is greater than 0.34). The diversity variable was constructed using demographic data provided by ‘World Fact book’ of the Central Intelligence Agency of the United States. The electricity consumption variable indicates consumption of electric power kilowatt-hour per capita. ‘Tertiary enrolment’ variable indicates percent of school enrolment in post-secondary educational programs (i.e., percent of students who successfully completed high school and enrolled in college level degree programs). The data on these variables were obtained from World bank database. The dummy variables ‘Anglophone’, ‘Francophone’, and ‘Hispanophone’ refer to groups of countries where main official or de facto languages are one of the following: English, French or Spanish. Data on remaining variables were collected from World Bank database, Federal Reserve Bank of St. Louis (FRED), and IMF database. The list of ‘Anglophone’ countries is obtained from the British Commonwealth secretariat website. The list of ‘Francophone’ and ‘Hispanophone’ countries were taken from ‘One World-Nation’s Online’ (www.nationsonline.com).

We expect all the regressors in the baseline model to have positive influence in per capita GDP. For the variant model, my a priori expectation is that the dummy variable ‘Anglophone’ will display statistically significant influence on the dependent variable. My intuition is partly shaped by Dahl (1985) and Dahl and Velupillai (2011) regarding differences in languages in terms of their expressions implying future events or actions, and partly by Chen (2013) regarding impact of language in shaping user’s economic behavior. According to Dahl (1985), English language has no inflectional marking for expressions regarding future events or actions whereas for French and Spanish languages such inflectional markers exist. According to Chen (2013) language impacts economic behaviors. We contend that since English language is less distinctive in referring to future events or actions (no inflectional future marker) than French and Spanish languages (both languages have inflectional future marker), speaker of English language will have difference in economic behavior relative to their French and Spanish counterparts which will, in turn, contribute to difference in per capita real GDP in Anglophone countries relative to other language cohorts.

DATA ANALYSIS AND RESULTS

Table 1 presents a list of Anglophone, Francophone, and Hispanophone countries. According to per capita GDP in 1997, top ten countries in terms of per capita GDP among the Anglophone countries were Brunei Darussalam ($75,654), Singapore ($45,863), United States ($43,166), Canada ($33,310), Ireland ($32,568), Australia ($32,013), United Kingdom ($30,614), Cyprus ($26,704), New Zealand ($26,152), and, the Bahamas ($22,971). Interestingly, none of these countries made it to the list of top ten Anglophone countries ranked in terms of growth in per capita GDP between 1997 and 2013. The top ten countries in terms of
per capita GDP growth during this period (1997-2013) are: India (per capita GDP growth rate 134.68%), Sri Lanka (120.76%), Sierra Leon (94.80%), Rwanda (93.42%), and Nigeria (92.78%), Bangladesh (91.28%), Zambia (82.52%), Ghana (80.26%), Mauritius (73.45%), and Singapore (72.16%). Interestingly, most of these fast growth countries are generally known as developing countries which are perhaps catching up with their industrially advanced counterparts as was predicted by Solow’s (1956) growth theory and was validated by several empirical studies on this topic (e.g., Barro, 2012; Baumol, 1986; Mankiw, Romer, and Weil, 1992).

Top ten Francophone countries in terms of per capita GDP in 1997 are: Canada ($33,310), Gabon ($2,545), Seychelles ($17,272), Algeria ($59,322), Equatorial Guinea ($7,037), Tunisia ($6,649), Republic of Congo ($4,618), Morocco ($4,177), Cote d’Ivoire ($3,131), and Vanuatu ($2,693). In terms of growth in per capita income for the period 1997 and 2013, top ten countries are: Equatorial Guinea (364%), Chad (95%), Rwanda (83%), Morocco (67%), Burkina Faso (64%), Tunisia (62%), Seychelles (44%), Algeria (38%), Djibouti (36%) and Mali (35%).

Top ten countries in terms of highest per capita GDP growth rate are: Bolivia (665.79%), Honduras (440.74%), Nicaragua (151.30%), Colombia (147%), Costa Rica (133.89%), Guatemala ($124.10%), Chile (123.33%), Dominican Republic (52.37%), Ecuador (51.61%), and Mexico (19.89). Here again we observe that half of these ten countries were not among the top ten Hispanophone countries measured by per capita income in 1997 which perhaps is an indication that countries once poor are recently experiencing faster growth creating potential for convergence of per capita income. Out of forty-four Anglophone countries, four countries’ experienced negative growth in per capita GDP between 1997 and 2013. These four countries are: the Bahamas (-1.97%), Barbados (-5.14%), Solomon Islands (-5.22%), and Brunei Darussalam (-8.17%). Out of twenty-six Francophone countries, four countries experienced negative growth in per capita GDP. These countries are: Madagascar (-2%), Togo (-5%), Gabon (-14%), Central African Republic (-31%). Out of sixteen Hispanophone countries four countries posted negative
### Table 1: Cohorts of Countries by Dominant Language

<table>
<thead>
<tr>
<th>Anglophone Countries</th>
<th>Francophone Countries</th>
<th>Hispanophones Countries</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antigua and Barbuda</td>
<td>Papua New</td>
<td>Algeria</td>
</tr>
<tr>
<td>Australia</td>
<td>Rwanda</td>
<td>Benin</td>
</tr>
<tr>
<td>Bahamas, The</td>
<td>Saint Lucia</td>
<td>Burkina Faso</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>Samoa</td>
<td>Burundi</td>
</tr>
<tr>
<td>Barbados</td>
<td>Seychelles</td>
<td>Cameroon</td>
</tr>
<tr>
<td>Belize</td>
<td>Sierra Leone</td>
<td>Canada</td>
</tr>
<tr>
<td>Botswana</td>
<td>Singapore</td>
<td>Central African</td>
</tr>
<tr>
<td>Brunei Darussalam</td>
<td>Solomon Islands</td>
<td>Chad</td>
</tr>
<tr>
<td>Cameroon</td>
<td>South Africa</td>
<td>Comoros</td>
</tr>
<tr>
<td>Canada</td>
<td>Sri Lanka</td>
<td>Congo, Republic of the</td>
</tr>
<tr>
<td>Cyprus</td>
<td>St Kitts and Nevis</td>
<td>Côte d'Ivoire</td>
</tr>
<tr>
<td>Fiji</td>
<td>St Vincent and</td>
<td>Democratic</td>
</tr>
<tr>
<td></td>
<td>Tonga</td>
<td>Equatorial Guinea</td>
</tr>
<tr>
<td></td>
<td>Trindia</td>
<td>France</td>
</tr>
<tr>
<td></td>
<td>Tuvalu</td>
<td>France Guiana</td>
</tr>
<tr>
<td>India</td>
<td>Uganda</td>
<td>French Polynesia</td>
</tr>
<tr>
<td>Ireland</td>
<td>United Kingdom</td>
<td>Gabon</td>
</tr>
<tr>
<td></td>
<td>United</td>
<td>Guadeloupe</td>
</tr>
<tr>
<td>Kenya</td>
<td>United States</td>
<td>Guinea</td>
</tr>
<tr>
<td>Kiribati</td>
<td>Vanuatu</td>
<td>Haiti</td>
</tr>
<tr>
<td>Lesotho</td>
<td>Zambia</td>
<td>Madagascar</td>
</tr>
<tr>
<td>Malawi</td>
<td></td>
<td>Mali</td>
</tr>
<tr>
<td>Malaysia</td>
<td></td>
<td>Martinique</td>
</tr>
<tr>
<td>Maldives</td>
<td></td>
<td>Mauritius</td>
</tr>
<tr>
<td>Malta</td>
<td></td>
<td>Monaco</td>
</tr>
<tr>
<td>Mauritius</td>
<td></td>
<td>Morocco</td>
</tr>
<tr>
<td>Mozambique</td>
<td></td>
<td>New Caledonia</td>
</tr>
<tr>
<td>Namibia</td>
<td></td>
<td>Niger</td>
</tr>
<tr>
<td>Nauru</td>
<td></td>
<td>Reunion</td>
</tr>
<tr>
<td>New Zealand</td>
<td></td>
<td>Rwanda</td>
</tr>
<tr>
<td>Nigeria</td>
<td></td>
<td>Senegal</td>
</tr>
<tr>
<td>Pakistan</td>
<td></td>
<td>Seychelles</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Togo</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tunisia</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Vanuatu</td>
</tr>
</tbody>
</table>

[Source: List of Anglophone countries is compiled from commonwealth Secretariat publication (http://thecommonwealth.org/member/- countries#sthash.78HRkmE.dpuf). Lists of Francophone and]
Hispanophone countries are collected from Nations Online (URL: www.nationsonline.org).

Growth in per capita GDP between 1997 and 2013. These countries are: Panama (-23.53%), Cuba (-25.64%), Paraguay (-33.49%), and Equatorial Guinea (-40.97%).

Table 2 presents the analysis of variance (ANOVA) including specification for Bonferroni pairwise comparisons. The ANOVA results reject the null hypothesis and confirm that there exists statistically significant difference between language groups (cohorts) mean GDP. Pairwise, difference in mean GDP of Anglophone countries are greater than that of Francophone and Hispanophone countries. Bonferroni corrected p-value is 0.012 for Anglophone and Francophone pair mean GDP comparison and p-value is 0.031 for the Anglophone and Hispanophone pair mean GDP comparison. We fail to reject the null hypothesis for the Francophone and Hispanophone cohorts as the difference in mean GDP of these two cohorts are not found to be statistically significant. We use generalized least square (GLS) model to correct for heteroskedasticity as diagnostic test (modified Wald test and log likelihood ratio test) results called for rejection of the null hypothesis of homoscedasticity.

**TABLE 2 ANOVA**

Summary of Mean GDP for three language cohorts

<table>
<thead>
<tr>
<th>Language group</th>
<th>Mean GDP</th>
<th>Std. Dev.</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anglophone countries</td>
<td>1.207e+10</td>
<td>9.477e+11</td>
<td>2,175</td>
</tr>
<tr>
<td>Francophone countries</td>
<td>5.048e+10</td>
<td>2.871e+11</td>
<td>1,242</td>
</tr>
<tr>
<td>Hispanophone countries</td>
<td>5.275e+10</td>
<td>1.363e+11</td>
<td>972</td>
</tr>
<tr>
<td>Total</td>
<td>8.578e+10</td>
<td>6.882e+11</td>
<td>4,389</td>
</tr>
</tbody>
</table>
Analysis of Variance

<table>
<thead>
<tr>
<th>Source</th>
<th>Sum of squares</th>
<th>Degree of freedom</th>
<th>Mean sum of squares</th>
<th>F-statistic</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between groups</td>
<td>5.2584e+24</td>
<td>2</td>
<td>2.6292e+24</td>
<td>5.56</td>
<td>0.0039</td>
</tr>
<tr>
<td>Within groups</td>
<td>2.0729e+27</td>
<td>4386</td>
<td>4.7261e+23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>2.0781e+27</td>
<td>4388</td>
<td>4.7359e+23</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Bartlett's test for equal variances: $\chi^2(2) = 3.9e+03$ Prob $\chi^2 = 0.000$

Comparison of GDP means across language cohorts (Bonferroni)

<table>
<thead>
<tr>
<th>Row mean – Column mean</th>
<th>Anglophone countries</th>
<th>Francophone countries</th>
</tr>
</thead>
<tbody>
<tr>
<td>Francophone countries</td>
<td>-7.0e+10 (0.012***)</td>
<td></td>
</tr>
<tr>
<td>Hispanophone countries</td>
<td>-6.8e+10 (0.031**)</td>
<td>2.3e+09 (1.00)</td>
</tr>
</tbody>
</table>

Note: P-values in parentheses are with Bonferroni corrections.

Table 3 summarizes results of the generalized least square model panel data regression. The estimated coefficient of the variable R&D share of GDP is positive and statistically significant at five percent level for both baseline model and the variant model with dummy variables indicating language cohorts. According to Table 3 results, one percent increase in R&D share will increase GDP growth rate by 0.02 percent in the base line model column, will increase GDP growth rate by 0.028 percent as shown in the variant model column.
Khan

TABLE 3 GLS REGRESSION OF GDP GROWTH RATE

<table>
<thead>
<tr>
<th>Dependent variable: GDP growth rate</th>
<th>GLS Baseline model</th>
<th>GLS-variant model with dummy variables for language cohorts</th>
</tr>
</thead>
<tbody>
<tr>
<td>R&amp;D share of GDP</td>
<td>0.019**</td>
<td>0.028***</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.010)</td>
</tr>
<tr>
<td>Trade Openness</td>
<td>0.028***</td>
<td>0.018**</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>Internet usage</td>
<td>0.003</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>Ethnic Diversity</td>
<td>0.010***</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>Electricity consumption</td>
<td>-0.005</td>
<td>-0.007</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>Tertiary enrollment</td>
<td>0.014**</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>Foreign investment inflow as percent of GDP</td>
<td>0.007***</td>
<td>0.007***</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Anglophone dummy</td>
<td>-</td>
<td>0.059***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.018)</td>
</tr>
<tr>
<td>Francophone dummy</td>
<td>-</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.001)</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.217**</td>
<td>-0.186**</td>
</tr>
<tr>
<td>No. of observations</td>
<td>575</td>
<td>575</td>
</tr>
<tr>
<td>Wald chi2(9)</td>
<td>95.36</td>
<td>113.44</td>
</tr>
<tr>
<td>Probability &gt; chi2</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Log likelihood</td>
<td>320.8683</td>
<td>325.6311</td>
</tr>
</tbody>
</table>

GLS standard errors in reported in parentheses. Statistical significance of the estimated coefficients is indicated by asterisks using conventional notions of 10% (*), 5% (**), and 1% (***)
levels. Binary variables Anglophone takes on value of 1 when a country’s primary official language is English and zero (0) otherwise. Binary variables Francophone, and Hispanophone are created in similar manner.

This result is mostly consistent with existing literature (e.g., Grossman and Helpman, 1990; Jones, 1995; Romer, 1990; Stokey, 1995). The estimated coefficients for the variable ‘trade openness’ are statistically significant at 1 percent level for both the base line and variant models. The results imply that a one percent increase in trade openness will accelerate GDP growth rate by 0.28 percent (per the baseline model) and 0.018 percent (per the variant model results). The estimated coefficients for internet usage are positive but turned out to be not
statistically significant. The estimated coefficient for the independent variable ‘ethnic diversity’ is found to be positive and statistically significant at one percent level in baseline model and positive but not statistically significant in the variant model with language cohort dummy variables. The existing literature is mixed about the impact of diversity on economic growth (e.g., Khan and Roy, 2011; Lazear, 1999; Watkins and Ferrara, 2005). Per obtained results in the baseline model, a one percent increase in ethnic diversity will boost GDP growth rate by .01 percent. Variable of per capita electricity consumption was not found statistically significant and the negative sign of the coefficients were counterintuitive. Tertiary enrolment variable’s estimated coefficients are positive for both specifications, and is statistically significant for the baseline model. According to the estimated coefficient values, a one percent increase in tertiary enrolment will increase GDP growth rate by 0.014 percent. Conforming to existing literature, foreign investment inflow is also a positive determinant of per capita GDP. Using a cross country dataset covering 1985 to 1995, Knack and Keeper (1995) report high associated expropriation risk with foreign direct investment (FDI) for recipient country. Despite being risky, FDI is a positive determinant of economic growth as found in several previous papers (e.g., Nair-Reichert and Weinhold, 2001; and Choe, 2003). In this paper, the estimated coefficient for the variable FDI share of GDP is statistically significant at 1 percent level. Per regression result, 1 percent increase in the variable foreign investment inflow as percent of GDP will increase per capita GDP by about 0.01 percent in both baseline model and variant model.

In the variant model, we use Hispanophone countries’ as baseline group and include Anglophone and Francophone cohort dummy variables as regressors. The estimated coefficient for Anglophone countries is positive and statistically significant at 1 percent level in the variant model. According to regression results, Anglophone countries’ per capita GDP growth rate is 0.06 percent more than of the baseline cohort (i.e. Hispanophone countries). We contend that that this difference in GDP growth across language cohorts is huge and calls for further empirical studies of these language cohort countries not only from the perspective of potential institutional differences in colonial legacies but also from a view point that difference in economic behaviors induced by the differences in FTR across languages which may influence social psyche differentially with respect to economic use of time. The varied sense of urgency in valuing time for its economic use can arguably lead to difference in economic behavior which can, in turn, contribute to difference in real GDP and economic growth. We believe, future empirical research efforts to extend literature on cross-country economic growth determinants should explore suitable proxy variables to explain economic growth differentials across countries due to institutional efficiency differentials and due to language induced economic behavior differentials.
CONCLUSION

In addition to exploring influence of selected macroeconomic variables on economic growth rates across countries, this study finds evidence that countries with different language cohorts differ in economic performance measured by real GDP and that this difference is statistically significant for pairwise comparison of means. Several studies cited in this paper identified variations in institutional legacies as a source of difference in economic growth across rates across countries. In this paper, we proposed an additional hypothesis that income differential across countries may also be influenced by differences in economic behaviors of agents shaped by difference in FTR across languages, or difference in ‘inflection’ or ‘non-inflection’ across languages in referring to future events or actions. In addition to be serving as proxies for institutional differentials, language-cohorts can also be a proxy variable measuring economic behavior differentials (and consequent productivity differentials) which ultimately contribute to difference in economic performance measured by real GDP across nations. Future research endeavors should be directed in studying potential influence of differences in linguistic structures regarding future time reference on differences in worker productivity and economic growth.

Dahl (1985) and Dahl and Velupillai (2011) analyzed many languages and identified English to be a language having no inflection in marking future events or actions, and identified French and Spanish as languages with inflection marking future events or actions. We contend that this difference in lingual expression (i.e., explicit vs. implicit use of verb indicating future tense) poses differential impact on users’ mindset in terms of intertemporal decision making (i.e., economic decisions involving present and future periods under some sort of tradeoff perspectives). One limitation of the paper is data availability. Although GDP data and data on most of the regressors used in this study are available for many developed countries since 1960, for most of the developing countries data is available only for more recent years since late 1990s or later years. Few countries’ yearly data on the desired variables were not available until after 1997. Therefore, due to data availability issue and missing data issue the regression could not capture longer time horizon and could not include all the countries supposed to be member of these three language cohorts. As more and more data covering longer time horizon becomes available covering more countries in the world, the dataset will be enriched and subsequent regression and other forms of statistical analyses will also be more robust. Future researcher can include more countries to the panel dataset by including German, Portuguese, and other language cohort countries and use new proxy variables to measure impact of different colonial institutional legacies and impact of different economic behaviors by agents induced by difference in colonial language FTR driven social psyches on economic performance measured by real GDP, labor productivity etc.
REFERENCES


PRESIDENTIAL ELECTIONS AND INDUSTRY STOCK RETURNS: A TEST OF MARKET EFFICIENCY

Patrick Gobran
Frank Bacon
Longwood University

ABSTRACT
The purpose of this study is to test the efficient market theory, and to see if individual stock industry returns can correctly predict the outcome of a Presidential Election. Specifically, will returns in industries that stand to benefit from the platform of a particular candidate show excess gains in the event-period either prior to or after Election Day? Is it possible to earn returns that are above normal, by having accurate polling data leading into an election? The link of election outcomes and stock market returns has been heavily studied and there is research that shows how certain market aspects can be seen to predict an election outcome. Likewise, we have seen how following a surprise election victory, certain industries have seen supra-normal returns for days and even weeks. According to the semi-strong form efficient markets theory proposed by Eugene Fama, one should not be able to achieve above average returns through the use of public information such as polls. All publically available information should already be factored into the stock price, as such meaning it is not possible to outperform the market, when adjusting for risk.

Key Words: Event Study, Finance, Market Efficiency, Elections, Excess Returns

BACKGROUND
The correlation between presidential elections and the stock market is a topic which has been heavily debated for some time. The creator of the Stock Traders Almanac – Yale Hirsch, pioneered the presidential cycle theory, explaining that the market will struggle in the first two years of a presidency, before generally being much more positive in the latter years of a term. This suggests that the market factors in the uncertainty which a new president can bring, and that it takes some time for any significant political changes to come about under a new administration.

Because of the close relationship between presidential election cycles and the stock market, it appears that that if you closely followed certain market indicators, then you could potentially predict presidential elections before the actual election day. This is somewhat true, and there is a great deal of ongoing research into this concept. According to InvesTech Research (2012) the stock market has been the most successful indicator of who will win the presidency over the past 100 years. Stock markets have predicted the winner over 90 percent of the time; with a market rise or fall on Election Day (prior to votes being tallied) being a major factor in this predictive power. (Kates Smith, 2016).

This predictive factor of the market may however not be predictive at all. Rather, this may be an example of market efficiency; a concept first published by Eugene Fama in 1969. Based on Fama’s theory of the efficiency of capital markets,
the current values of a stock reflect all available information. Fama describes that there are three possible levels of market efficiency: strong form, semi-strong form, and weak form. (Fama, 1970, 1997).

The notion that the market can predict a presidential election winner strikes at the heart of the efficient market hypothesis from the finance literature. Specifically, this suggests that the market is semi-strong form efficient. Based on this type of analysis, all public polls and election data will already be reflected in the price of a stock, while “private” polls or information on the candidates might allow traders to gain a leg up and earn above average returns (an example of “insider trading” to a certain extent). Rather than predicting the winner, the market is merely reacting to the most possible or likely outcome of the election, before Election Day even comes. (Foster, 2012).

Additionally, there would likely be different market reactions based on polling leading up to an election. It follows that different market reactions should be observed depending on whether the Republican party or the Democratic party takes the White House. Likewise, industries’ which lean towards either party should exhibit either positive or negative stock returns accordingly.

Recently, we observed the impact a surprise election result can have on the market, with both the British Brexit Referendum and the 2016 United States Presidential Election. In both cases the status-quo candidate/choice was upset by a narrow margin, which went against the vast majority of public polling data. In both cases, the outcome surprised analysts and traders around the world. The surprise election result of the Brexit sent shockwaves throughout the markets, embracing the notion that traders dislike the uncertainty this created. The immediate drop in markets after the results, as well as the fluctuations observed as the votes were tallied shows how fast financial markets can respond to relevant information. Likewise, with the results of the 2016 Presidential Election, stock markets took huge losses late into the night as President-Elect Trump pulled ahead, before rebounding sharply the next day as many industries saw potential benefits with his administration. In both cases, the stock market had already “priced-in” the most likely outcome, and the “surprise” deviance from the status-quo led to great fluctuations in global equity markets. (Yan, Riley & Egan, 2016).

PROBLEM & PURPOSE
How do Democratic and Republican leaning industries stock returns react to Presidential Election outcomes? Is it possible for the market to predict an election outcome? When looking at specific industries over the course of a Presidential Election cycle, can we see excess returns previous to Election Day?

The purpose of this research is to test market efficiency by examining how the risk adjusted rate of return for stocks that lean specifically towards a Democratic or Republican Presidential candidate react to election outcomes. Two 15 firm samples from industries that are defined as favoring the policies of the incumbent candidate – President Bush – in the 2004 Election will be randomly selected and tested using the standard risk adjusted event study methodology found in finance literature. Two different 15 firm samples from industries that are defined as favoring the polices of the incumbent candidate – President Obama – in the 2012 Election
will also be selected and tested. Lastly, two 15 firm samples from industries which would benefit from the policies of President Trump (2016) were selected. Due to Presidents Bush, Obama, and Trump winning the 2004, 2012, and 2016 Elections (respectively), we would expect a positive reaction from these industries, as they stood to do better under the winning presidents, than under their opponent’s presidencies. This event study tests whether the results of these Presidential Election exhibit weak, strong, or semi-strong market efficiency theory as defined by the efficient market hypothesis in the finance literature.

LITERATURE REVIEW

For the purpose of this research, we will look at stocks in four Republican and two Democratic leaning industries. These sectors/industries will be divided based on research conducted by a Gallup poll, basing party affiliation on political donations, and by breaking down the potential effect of a candidate’s platforms on certain industries. There is some difficulty with dividing industries, as within an industry such as utilities, solar and alternative energy would be Democratic leaning, while coal, oil and gas would be Republican leaning. Further complicating the matter is the fact that some industries seemingly give to the incumbent party, regardless of historic party affiliation. (Newport, 2013).

As mentioned, some industries may donate to both parties, thus ensuring potential favorable treatment on contentious policy issues. Along with this, some industries may donate to down ballot candidates for one party, while supporting an opposing Presidential candidate. For this reason, we will be analyzing industries which heavily favored a certain candidate. For 2004, the two Republican leaning industries will be Financial Services, and Aerospace/Defense. Financial Services is a strong choice, as President Bush supported a policy of financial deregulation, and would be considered beneficial to Wall Street. Giving from the Financial Services industry has at times been 7 to 1 in favor of the Republican Party, yet it should be noted that in 2004, large financial firms were some of the biggest individual donors to both candidates. President Bush, who was in favor of the Iraq War, proposing either maintaining or increasing the defense budget, while Senator Kerry wanted to curb or reduce American involvement in the Iraq War, deferring responsibility to NATO partners. (Gimpel, Lee & Parrott, 2012; Demasters, 2012)

For the 2012 Election, healthcare and alternative energy industries stood to most heavily affected. President Obama and his Affordable Care Act would have huge impacts on healthcare firms, while the election of Romney could lead to future uncertainty for the healthcare industry. Likewise, President Obama made alternative energy a matter of high importance, providing government funding and support to this industry. Romney, like many Republicans catered more towards traditional fossil fuel and coal energy providers, as such, his election would have negatively affected alternative energy firms. (Newport, 2013; “US Presidential Election 2012”)

In 2016 Election, the financial services industry and what would be considered the “traditional” energy industry will be analyzed. Due to the proposals from President Trump on the campaign trail to repeal Dodd-Frank, and to cut regulation in the financial services industry, large domestic banks had huge upside potential. Likewise, Trump made it a cornerstone of his platform to “bring-back”
coal jobs, and to cut a great deal of environmental regulation that came around during President Obama’s years in the White House. Unlike the 2004 and 2012 samples, where we had an incumbent President, who was re-elected, the 2016 Election saw a first-term President with many policy “unknowns” get elected.

The reasoning behind choosing these elections was their tightly contested natures. The 2004 Election was a very close race between the incumbent Republican candidate, President Bush, and the Democratic candidate, Senator John Kerry, with the popular vote within 2 percentage points. Due to the close nature of this race, polling compiled by Real Clear Politics in the 30 days leading up to the election had a national average lead for President Bush of 3.04%. However, the last nationally published poll, on the day prior to election day had Senator Kerry up by 1 point. By this logic, the market could look to the polling data, and get a good - yet not 100% - idea of who would win the election. (Real Clear Politics, 2004)

In the case of the 2012 Election, which proved to be even closer in the weeks leading to Election day, Mitt Romney offered a serious challenge to incumbent President Obama. Much like the 2004 election, polls in the 30 days prior to election day were inconsistent, with neither candidate having a large advantage. The national polling average in the 30 days prior to the election saw President Obama leading by 0.4%; a number that could be considered within the margin of error. This parlayed into election day, with the candidates finishing within 4 percentage points of each other, in a race that was too close to call for many hours. (Real Clear Politics, 2012)

When looking at the 2016 Election, polls routinely had Hillary Clinton, the Democratic candidate for President, in a lead of up to 14% within 30 days of the Election. Of the 60 polls released in the 30 days leading up to the Election, only a handful had President Trump coming out victorious, with an average lead of over 5% for Clinton. This makes the 2016 Election different than the previous two samples, and should reflect the markets being surprised by the result. With the 2016 Election, stock markets were rocked as the early results came in, with Dow futures trading down over a full percent, before markets reacted to the potential upside of a Trump presidency, rallying a great deal before opening the next day. (Real Clear Politics, 2016).

According to the efficient market hypothesis, there are three forms of market efficiency: strong form efficiency, semi-strong form efficiency, and weak form efficiency. Strong form efficiency is the highest level of efficiency. It states that all information is accurately and currently reflected in the market prices of securities. This includes public and private or “insider” information. In a strong form efficient market, even investors who have insider knowledge will not be able to earn abnormally large profits. Weak form market efficiency states that current stock prices reflect all past information. This then means that a trader cannot use analysis of past stock prices to accurately forecast future stock prices, as this information is already reflected in the current market price. Lastly, and located between strong form and weak form efficiency, is semi-strong form market efficiency. Semi-strong form efficiency states that the current stock price reflects all publically available information, including past information. However, because only publically available
information is reflected in the stock price, insider or private information could be used by a trader to make abnormally large (and illegal) profits. (Fama, 1970,1997).

In reality, there is a great deal of evidence suggesting that markets are semi-strong form efficient; insiders can make abnormal returns by acting on information before it is made public, and even once this information is made public, there can be some lag time before the price fully reflects all available information. This lag can vary in length depending both on the market, the individual security, and the way in which information is shared. This lag has decreased in recent years due to technology, and is less of a concern now than in the past. This lag is however reason for us to believe that the market is not strong form efficient. (Fama, 1970, 1997).

Market efficiency is affected by a number of factors, including: the number of participants (firms, investors, traders, etc…), transaction costs on trades, and the flow of information. The more participants, the more efficient the market will be, as there are more people driving the market forward, all for their own best interest. As more reliable and accurate information becomes available, and as the speed with which this information flows increases, the market becomes more efficient. With low costs on trades, a market will be more efficient, as the cost of performing trades will not hinder investors from reacting to price fluctuations. (Fama, 1970,1997; Bacon & Labbs, 2013; Bacon & Greis, 2008; Bacon & Kinsler, 2008).

With the efficient market hypothesis, election results should be reflected in prices immediately and investors should not expect to earn a return higher than the normal rate of return, after the close of polls. After-hours, futures, and before-the-bell trades will likely reflect the outcome of the election in real time, as the polls close and results come in. However, because of the heavy publicizing of polls, and potential public sentiment (positive or negative) towards electoral candidates, the market may in fact be responding to information in the lead up to an election. That is, “pricing-in” information before the actual Election Day. The market will adjust to public polls before an investor has time to trade, even if they are aware of all the latest information (Foster, 2012).

Do presidential elections impact stock returns? If so, what level of market efficiency do we observe? More specifically, are there certain industries returns which are heavily impacted and can seemingly predict an election outcome before Election Day? To answer these questions, this study will analyze the risk adjusted stock returns of sample stock prices of two Republican & to Democratic leaning industries in the 30 days before and after Election Day (2004 & 2012). The analysis seeks to determine if the outcome of an election enables investors to earn an above normal risk adjusted return. Is it possible an investor to “beat” the market relying solely on public information in the form of pre-election polls? Is it possible to predict an election based on the stock market? (Fama, 1970,1997; Bacon & Labbs, 2013; Bacon & Greis, 2008; Bacon & Kinsler, 2008).

**METHODOLOGY**

This study will use the standard risk adjusted event study methodology in the finance literature. The Election Day will be Day 0. The required historical financial data, such as the movement of S&P500 during the event study period will also be obtained from Yahoo! Finance.
In order to test semi-strong market efficiency with respect to Presidential Election outcomes, this study proposes the following hypotheses:

**H1**: The risk adjusted return of the stock prices of the sample firms are not significantly affected by the results of the presidential election, around election day as defined by the event period.

**H1**: The risk adjusted return of the stock prices of the sample firms are not significantly affected by the results of the presidential election, around election day as defined by the event period.

The firms that will make up the 2004 Republican sample will be randomly selected from the Financial Services, and Aerospace/Defense industries, while the firms that will make up the 2012 Democratic sample will be randomly selected from the Healthcare, and Alternative Energy industries. Finally, the firms which make up the 2016 Republican sample will be randomly selected from the Financial Services, and Traditional Energy industries.

**QUANTITATIVE TESTS AND RESULTS**

The results of graphing the CAER over time for both 2004 Republican industries are not in line with our expectations, while the results of graphing the CAER over time for Healthcare in 2012 (as seen below) and Financial Services in 2016 more closely align with our predictions. A 2 sample t-test was performed, on both AER and CAER against a return of 0 (where expected returns would be equal to actual returns).

In 2004, for both industries we saw negative CAER’s almost right from the beginning, with small to medium sized bumps in the days around the election, before the CAER declined again until the end of the event-study period. This did not line up with our hypothesis.

Conversely, both the Healthcare and Alternative Energy industries saw positive CAER’s in 2012, The fluctuations in return between 2004 and 2012, against the expectation of similar returns will be explained in the Findings section.

Lastly, for the 2016 election, significant results for Financial Services were seen when analyzing CAER.
CONCLUSION

Our findings were in a sense surprising, yet at the same time very much in line with what we would expect if we assume that the semi-strong form model of market efficiency is correct. Because polling data is made public, and as a result is acted on by rational investors, the predicted outcome (which in the 2004 Election was for President Bush to win and in 2012 was for President Obama to be victorious) was already “priced-in” to the stocks. This correlation between polling data and industry stock returns can best be illustrated when we look at AER’s in the Healthcare industry against polling data for the 2012 Election. As we would expect, on days when polls were released that had President Obama in the lead, Healthcare industry stocks often increased. This became more the case as we moved closer to Election day, seeing a correlation of 0.34 over the entire course of the 30 days prior to the Election. This correlation includes polls which may not have been taken as seriously, due to perceived bias, while additionally averaging polling results on days when multiple polls were released. In terms of the 2016 Election, the expected outcome was not the actual result, and we saw excess returns for industries which served to benefit from this surprise; again, this makes sense.

However, due to all the elections being a fairly close race in terms of polling, the markets would be thought to have some reaction upon the Election results becoming official. This is very prevalent when looking at the 2012 Election, and healthcare returns the day after the Election (Day 1), as the AER for the 15 healthcare firms shoots up significantly. This was due to a confirmation of polling data, that then allowed investors to act with some confidence. In 2004, President Bush did have an average lead of over 3% in the month leading up to Election Day, meaning that this was in no way a surprise victory. Because of this, and the fact Bush was the incumbent President, his policies and their potential market impacts were already well known. As such, policies were priced in over the months leading up to the election, and other macro and micro-economic events were larger drivers of industry specific stock prices.

In terms of the 2016 Election, the prediction of a Clinton victory, and the continuation of Obama Administration policies was expected, and priced into stocks. When President Trump won, stocks in industries which could benefit under his
presidency saw excess returns. However, unlike the previous elections (2004 & 2012), these policies would not be enacted until after Inauguration in January, thus leading to some potential skew with returns in the 30 days after Election day, as Obama policies would still have an impact on the selected industries in the short term. This is very likely to have been the case with the Traditional Energy industry, which saw some excess returns following the Election, before seeing a steep drop. The firms selected for the Traditional Energy sample include coal companies as well as firms which would stand to benefit from the approval and completion of projects such as the Keystone XL and Dakota Access Pipelines. With President Obama still being in office, these firms likely saw a temporary over-reaction in stock price following Election Day, then followed by a correction. Additionally, many of the firms in the sample are highly dependent on commodity prices, including oil, natural gas, and coal; fluctuations in the prices and/or demand for these commodities would serve to impact stock prices to a greater magnitude than the predicted future policies of a President.

Negative cumulative excess returns were seen for both 2004 Republican leaning industries, meaning that in the event study window of the election, their returns were actually lower than those of the market once they were adjusted for the risk. There are many reasons why this could happen, with some potential issues being the increasingly negative public sentiment towards the Iraq War (potentially hurting Aerospace/Defense firms), the relatively weak US Dollar, and the Federal Reserve deciding to increase interest rates. Additionally, the government of Japan announced an intent to cut military spending by 3% annually in mid-December, and due to the close military alliance between Japan and the United States, this cut in military spending would translate into less sales for American Defense/Aerospace firms. Conversely, for the Aerospace/Defense industry, the bankruptcy battles of United Airlines and US Airways throughout 2004, as well as the pending launch of the Airbus A380 could have caused Boeing, Lockheed Martin, and Embraer to all see their stock perform worse than the market. To eliminate the interference of non-defense sectors affecting overall firm stock performance (and somewhat skewing the results), pure-play defense firms would have to be used for the 15 firm sample. The pure-play concept is outlined in the finance literature, however due to the nature of US Government defense contracting, the elimination of any firm that conducted non-defense operations would leave a very small number of specialized publically traded firms from which to draw a sample.

Other potential reasoning for the negative CAER over the event period could include negative earnings surprises, bad publicity, or even policy decisions or elections outside of the United States. The weak US Dollar could have been a reason for the decline in financial services firms over the event period. However, a correlation test over the entire event study period comparing the average daily return of the financial services sample and the Trade Weighted US Dollar Index (Major Currencies) revealed a correlation of -0.11; in other words, this was not the reasoning behind these negative excess returns. A more likely issue could have been uncertainty over future rate hikes by the Federal Reserve.
It seems to be that Presidential Elections are not the most important thing in the market, and when looking at a status-quo or incumbent President, they can have very little impact on specific industries at all. However, when we are looking at the 2012 Election, we see positive CAER’s for both industry samples; part of this could be attributed to the closer polling for this election when compared to 2004. Healthcare saw a bump over the event period as the prospects of an Obama victory became larger and larger, with a large spike immediately after Election Day, as uncertainty over the future of the Affordable Care Act was erased. In the days following the post-Election bump however, industry stocks saw some negative excess returns. Some of this could be attributed to a correction in stock prices, in addition to the November 15th refusal by Texas Governor Rick Perry to operate a healthcare exchange in Texas. (Morton, 2012).

When we look at the alternative energy industry sample, we see substantial positive excess returns in the weeks following Election day. There are numerous reasons for this. From 2001-2009, there was a correlation of 0.86 between the per barrel price of Cushing, Oklahoma West Texas Intermediate Crude Oil and the Wilder Hill New Energy Global Index (NEX). This correlation is also seen over the event study period, yet to a lesser effect. Over the event study period, there was a 0.35 correlation between the daily change in per barrel price of Cushing, Oklahoma West Texas Intermediate Crude Oil and the average alternative energy sample industry return. This partially explains the bump in CAER following the re-election of President Obama. Another reason for this increase, was the high-volume of positive press about the alternative energy industry that was being published at the time. Germany had enacted green energy legislation over the past decade, and numerous reports about its effectiveness were being published. This would certainly have led to investor excitement, and as a result could explain some of the post-election bump. (Roen, 2015)

Based on the findings from the 2004 and 2012 Elections, the semi-strong form efficiency model seems to be supported by the evidence.

For the 2016 Election, the fluctuations in returns for Traditional Energy Stocks are briefly mentioned previously, while the returns for the Financial Services sector seem to line up with what we would expect. While we do see positive CAER’s prior to Election Day, this can be attributed partially to the large hit taken by the Financial Services industry after the results of the British Referendum or “Brexit” vote. This was during the event study period, and as a result we likely saw underpriced banking stocks. Additionally, while Clinton was expected to be tougher on large banks than President Trump, she also had a close relationship with some of the largest financial institutions in the country, previously giving paid speeches. Her husband, President Bill Clinton, had a large hand in deregulation, and would have been considered to be friendly to the industry. As a result of this, Financial Services firms were seeing excess returns even before Election day. Compounding this, times of uncertainty/volatility, lead to a higher volume of trades for investment banks, which could explain some of the excess returns prior to Election day.

The decrease in CAER’s for Traditional Energy stocks can be partially attributed to the fact that President Trump would not take office for a few more
months (as mentioned previously), in addition to the fact that while domestic policies may become more friendly to Traditional Energy, foreign policy could become harsher. In Canada, a Federal Carbon Tax program was first proposed in October, just prior to the US Election, before becoming formalized in early December (coinciding with a drop in CAER). Due to the fact that many of the companies in the Traditional Energy sector do business in Canada, this would have an effect. Additionally, fluctuations in commodity prices could account for decreases in the excess returns for Traditional Energy firms, as a decrease in the price of natural gas or coal could hurt many firms in the industry. Similar to the correlation between oil prices and Alternative Energy industry returns in the 2012 Election, we saw a 0.38 correlation between oil prices and Traditional Energy Stocks, with some fluctuations in the price of oil, along with major OPEC announcements accounting for a large amount of the negative excess returns seen around Day 20. (Austen, 2016)

Taking all of these findings into account, perhaps the only way you could beat the market in an election would be if you had private polling data that happened to be more accurate or reliable than what was publically available. If an investor had access to accurate, private data for the 2016 Election, they would have been able to earn above average returns by investing in portfolio that would stand to benefit from the policies of President Trump. If an investor purchased the 30 firms used for the sample, prior to Election Day (Day -1), they would have earned over 8% excess returns in just 30 days. It seems unreasonable however, for this to be possible, as the amount of resources which would be needed to have accurate data would likely outweigh any potential gains on stocks to be negligible. With this, the investor would also be taking additional risk that even with predicting the correct winner of an election, the stocks of an industry may not react as expected. What the research does seem to find evidence of however, is a market that is semi-strong form efficient.

REFERENCES


INTERNET TROLLING IN SOCIAL NETWORKING SITES: A PRELIMINARY INVESTIGATION OF UNDERGRADUATE STUDENT VICTIMIZATION

Carl J. Case
Darwin L. King
St. Bonaventure University

ABSTRACT

Internet trolls, often equated with online harassment, have gained public notoriety of late. As a result, this preliminary study was undertaken to explore the extent of trolling victimization of undergraduate business students, those that will be responsible for business communication in the future. Findings suggest that although there are five primary social networking sites utilized by students, nearly half of students had been trolled within the past six months and each student was trolled more than one time per month. Twitter users accounted for the largest percentage of students that were trolled while Snapchat and Instagram users had the highest incidence per victim. There were, however, mixed results with regard to the relationship between online social media usage time and trolling. Results imply that proactive and reactive measures may need to be implemented to minimize activity and assist victims in adequately dealing with the pressures and negative issues such as depression that may be associated with being trolled.

Key words: Internet trolling, social networking, YouTube, AACSB study

INTRODUCTION

Social media usage has seen incredible growth since 2010. A survey of more than 800 small business owners, for example, found that from 2010 to 2013, the percentage of businesses using LinkedIn increased from 37% to 57% (Ickert and McCracken, 2013). Facebook increased from 33% to 50%, Twitter increased from 12% to 26%, Google+ increased from 0% to 24%, YouTube increased from 0% to 19%, Pinterest increased from 0% to 6%, and Instagram increased from 0% to 2%. By August 2016, Twitter had 310 million monthly active users and LinkedIn had 106 million (Fortune, 2016). Moreover, in 2013, only 24% of US smartphone users age 18-24, 5% age 25-34 (millennials), and 2% age 35+ used the Snapchat app (Lella, 2016). By 2016, the user base had grown dramatically to 69% of smartphone users age 18-24, 38% age 25-34, and 14% age 35+. Interestingly, in 2014, 70% or 28 million online individuals age 51-70 had a Facebook account as compared to 88% of millennials (Carstensen, 2016; Wallis, 2014).
In concert with the increased usage of social media has been the increase in related challenges. The Internet Security Threat Report, for instance, indicates that social media remains problematic with regard to scammers leveraging the trust users have in their own social circles to spread scams, fake links, fake apps, fake plugins, and phishing (Symantec, 2016). In 2013, only 2% of social media threats required end-users to propagate them (Symantec, 2015). By 2014, this increased to 70%. In terms of social networking sites, Facebook, Twitter, and Pinterest users have been vulnerable to lifejacking and comment jacking that entices the user to click on a "continue" or "verification" button that in reality is a "like" button. Instagram scammers, on the other hand, mimic offers employed by legitimate corporate users. And, Snapchat, a self-destructing image and video platform, has been compromised by passwords obtained by other breached sites (in which users utilize the same password).

Another challenge that is emerging and gaining media attention is that of the Internet troll. An Internet troll is a person "who sows discord on the Internet by starting arguments or upsetting people, by posting inflammatory, extraneous, or off-topic messages in an online community (such as a newsgroup, forum, chat room, or blog) with the deliberate intent of provoking readers into an emotional response or of otherwise disrupting normal on-topic discussion, often for their own amusement" (wikipedia, 2016). Two online studies totaling over 1,200 respondents found that Internet users that describe themselves as trolls scored extremely high in the dark tetrad of personality traits: psychopathy, Machiavellianism, and, especially, sadism (Buckels, Trapnell, and Paulhus, 2014).

Trolling has been equated with online harassment. Troll behavior can range from pranks to harassment to violent threats and may include doxxing, the publishing of personal data such as Social Security numbers and bank accounts (Stein, 2016). In 2011, trolls visited Facebook memorial pages of recently deceased users to mock their deaths. In 2016, a New York Times editor quit Twitter after receiving anti-Semitic messages and a troll was permanently banned from Twitter after tweeting his 300,000 followers to harass the 2016 film Ghostbusters co-star Leslie Jones.

In fact, a Pew Research Center survey found that 73% of adult Internet users have seen someone harassed online and 40% report being personally harassed online (Duggan, 2014). For those personally harassed, 27% were called offensive names, 22% had someone try to purposefully embarrass them, 8% had been physically threatened, 8% had been stalked, 7% had been harassed for a sustained period, and 6% had been sexually harassed. Of note is that women were more likely to be harassed with 26% of women indicating being stalked online, as compared to 7% of men.

A 2016 Norton study further found that 76% of Australian women under the age of 30 had experienced abuse or harassment online (Griffith, 2016; Norton 2016).
The respondents indicated feelings of anger (36%), frustration (32%), irritation (32%), anxiousness (30%), depression (22%), helplessness (21%), fright (16%), and considered suicide (5%).

Given the recent media attention about trolling, this preliminary study was undertaken to better understand social media usage and trolling incidence among undergraduate students. The focus of this research is the business student population, in other words, those that will drive business communication in the future.

This study examines several questions. How active are students in social media today? What are the current primary social media sites utilized by undergraduates and is this usage similar to that of the Fortune 500 or AACSB accredited schools of business? What is the level of student victimization by Internet trolls? Have undergraduates seen others trolled online? Finally, is there a relationship between the time spent online at a social media site and the incidence of trolling? Results are important in determining if there is a need for reactive and/or proactive education with respect to trolling.

PREVIOUS RESEARCH
Researchers have examined several facets related to social networking. The authors conducted two studies to determine social media use by the top accredited business schools, the AACSB institutions, and by the largest firms, the Fortune 500. Additional studies have examined adoption, user satisfaction, and academic performance.

The two authors' studies were used to describe both the prevalence and types of social networking technologies employed in education and business. The first study reviewed the 647 AACSB accredited schools of business (Case and King, 2014). The authors found the most common social networking technologies utilized by institutions include Facebook (65% of schools), Twitter (57% of schools), YouTube (42% of schools), and LinkedIn (32% of schools). The least utilized technologies include blogs (24% of schools), Flickr (17% of schools), and other (15.8% of schools). "Other" included 36 different programs such as Daily Motion, Delicious, FourSquare, Google+, iTunes, Viadeo, Vimeo, and Weibo. The second usage study analyzed the Fortune 500 utilization of social networking technologies (Case, Ryan, and King, 2015). The most common technology was LinkedIn, with 97% of the firms implementing the technology. The other highly implemented products were Twitter (76% of firms), Facebook (74% of firms), YouTube (67% of firms), and blogs (53% of firms). The least commonly utilized technologies include Google+ (17% of firms), Pinterest (8% of firms), Flickr (6% of firms), and Instagram (3% of firms). "Other" social networking technologies included Forum, Foursquare, iTunes, Mobile Alerts, Slideshare, Spiceworks, StockTwits, Tumblr, and Viggie.
Relevant to adoption, researchers found a range of technological, organizational, and environmental factors that can impact the social media adoption decisions in Australian local government organizations (Sharif, Trishani, and Davidson, 2015). Technology factors include perceived benefits, perceived risk (fear that negative comments could be aired publicly by unauthorized parties), and compatibility with existing hardware. Organization factors include management drive, social media policies, and degree of formalization. Environment factors include community demand, bandwagon pressure, and faddishness.

Another study examined user satisfaction with social networking, in particular, Facebook (Wang and Li, 2015). Results suggested that competence and benevolence trust beliefs positively influence relatedness need satisfaction. Moreover, relatedness need satisfaction and relatedness need level significantly influence users' satisfaction with social networking sites. In other words, individuals feel more satisfied when their need for relatedness is satisfied (e.g., when they feel that others are friendly and feel that others like them).

With regards to academic performance, research has indicated that students that performed higher on the personality traits of agreeableness and conscientiousness were less likely to be associated with emotional/physical problems related to social network use (Glass, Li, and Pan, 2014). In addition, Chinese students that reported spending excessive time on social networking sites and those that believe their academic performance suffers because of social network use are more likely to have lower academic performance. Interestingly, Hall found that individuals spend only 3.5% of their time on social media interacting with others by commenting or chatting while the remainder of their time is spent browsing (2016).

**RESEARCH DESIGN**

This study employs a survey research design. The research was conducted at a private, northeastern U.S. university. A Student Internet Troll instrument was developed by the authors and administered to undergraduate students enrolled in a School of Business course. The courses included a variety of subjects such as Business Information Systems, Introduction to Financial Accounting, Introduction to Managerial Accounting, Macroeconomics, and Business Policy. A convenience sample of class sections and faculty members was selected. The surveys were collected during September of 2016.

The survey instrument was utilized to collect student demographic data such as gender and academic class. In addition, the survey examined student Internet behavior regarding online social media sites. Students were asked to estimate the average number of minutes spent daily on fourteen social media sites and list any other social networking sites used by the student. Moreover, students were prompted to estimate the number of times that he/she had been trolled on each site during the past six months and the number of times that he/she had seen others
trolled on each site during the past six months. Results were summarized by social media site and correlations were calculated to determine potential relationships between site usage minutes and the quantity of trolling incidences.

All surveys were anonymous and completed in an academic classroom. The response rate was 100 percent. Students were also informed that results would have no effect on their course grade.

RESULTS

A sample of 239 usable surveys was obtained. Table 1 indicates that, overall, 64% of the respondents were male and 36% were female. These percentages were fairly consistent with the study university's School of Business student population.

TABLE 1
Gender Response Rate

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>64%</td>
</tr>
<tr>
<td>Female</td>
<td>36%</td>
</tr>
<tr>
<td>Count</td>
<td>239</td>
</tr>
</tbody>
</table>

The response rate by academic class was relatively equally distributed with a slightly higher percentage of sophomores participating. Table 2 illustrates that 12% of respondents were freshmen, 35% were sophomores, 24% were juniors, and 29% were seniors.

TABLE 2
Academic Class Response Rate

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freshmen</td>
<td>12%</td>
</tr>
<tr>
<td>Sophomore</td>
<td>35%</td>
</tr>
<tr>
<td>Junior</td>
<td>24%</td>
</tr>
<tr>
<td>Senior</td>
<td>29%</td>
</tr>
</tbody>
</table>

Responses were first examined with regard to activity level per social media site. Although 14 sites were provided on the survey instrument, each respondent was prompted to list any “other” social media sites that he/she utilizes. The “other” sites named included WhatsApp, Barstool, Wall Street Oasis, Tinder, and VSCO. Table 3 shows that there are five sites that are used by most students. These include Snapchat (95% of students), Instagram (88% of students), Facebook (81% of students), Twitter (76% of students), and YouTube (67% of students). LinkedIn and Pinterest are used by considerably less students, 24% and 20%, respectively. The
remaining sites were uncommonly used by undergraduates. These least utilized social media sites are Google+ (9% of students), Reddit (6% of students), Tumblr (3% of students), YikYak (3% of students), other (3% of students), 4chan (1% of students), 8chan (0% of students), and Voat (0% of students). In terms of usage, users of Twitter spend the most minutes per day (44 minutes). Moreover, users of Instagram spend 43 minutes per day, Snapchat spend 42 minutes per day, YouTube spend 39 minutes per day, other spend 36 minutes per day, Facebook spend 32 minutes per day, Pinterest spend 30 minutes per day, Reddit spend 30 minutes per day, and Google+ spend 24 minutes per day using the given social media site. The least active users include Tumblr (17 minutes), LinkedIn (12 minutes), YikYak (10 minutes), 4chan (10 minutes), 8chan (no users), and Voat (no users). Overall, all respondents indicated using at least one social media site with the average undergraduate spending 176 minutes (nearly 3 hours) per day on social media sites.

Relative to trolling, the highest percentage of site users being trolled includes YikYak (63% of users) and Twitter (32% of users). Lesser trolled sites include Instagram (19% of users), Facebook (14% of users), other (14% of users), and Snapchat (13% of users). The least trolled users include YouTube (7% of users), Reddit (7% of users), Google+ (5% of users), and LinkedIn (2% of users). All other users reported not being trolled. Finally, users of each site were prompted to estimate the incidence of others that they have seen being trolled on the site. The social media sites that other individuals were most actively trolled include 4chan (100% of respondents), YikYak (75% of respondents), Twitter (67% of respondents), Facebook (64% of respondents), Reddit (64% of respondents), Tumblr (57% of respondents), Instagram (49% of respondents), and other (43% of respondents). The least trolled include YouTube (31% of respondents), Snapchat (20% of respondents), Google+ (14% of respondents), Pinterest (4% of respondents), LinkedIn (2% of respondents), 8chan (no users), and Voat (no users). Overall, 41% of students reported being trolled and 73% indicate noticing others being trolled at least once during the past six months.
### TABLE 3
**Social Media Activity and Trolling Percentages**

<table>
<thead>
<tr>
<th>Social Media Site</th>
<th>Percent of Students</th>
<th>Minutes Per Day</th>
<th>Percent of Students Trolled</th>
<th>Percent of Others Trolled</th>
</tr>
</thead>
<tbody>
<tr>
<td>Snapchat</td>
<td>95%</td>
<td>42</td>
<td>13%</td>
<td>20%</td>
</tr>
<tr>
<td>Instagram</td>
<td>88%</td>
<td>43</td>
<td>19%</td>
<td>49%</td>
</tr>
<tr>
<td>Facebook</td>
<td>81%</td>
<td>32</td>
<td>14%</td>
<td>64%</td>
</tr>
<tr>
<td>Twitter</td>
<td>76%</td>
<td>44</td>
<td>32%</td>
<td>67%</td>
</tr>
<tr>
<td>YouTube</td>
<td>67%</td>
<td>39</td>
<td>7%</td>
<td>31%</td>
</tr>
<tr>
<td>LinkedIn</td>
<td>24%</td>
<td>12</td>
<td>2%</td>
<td>2%</td>
</tr>
<tr>
<td>Pinterest</td>
<td>20%</td>
<td>30</td>
<td>0%</td>
<td>4%</td>
</tr>
<tr>
<td>Google+</td>
<td>9%</td>
<td>24</td>
<td>5%</td>
<td>14%</td>
</tr>
<tr>
<td>Reddit</td>
<td>6%</td>
<td>30</td>
<td>7%</td>
<td>64%</td>
</tr>
<tr>
<td>Tumblr</td>
<td>3%</td>
<td>17</td>
<td>0%</td>
<td>57%</td>
</tr>
<tr>
<td>Yik Yak</td>
<td>3%</td>
<td>10</td>
<td>63%</td>
<td>75%</td>
</tr>
<tr>
<td>Other</td>
<td>3%</td>
<td>36</td>
<td>14%</td>
<td>43%</td>
</tr>
<tr>
<td>4chan</td>
<td>1%</td>
<td>10</td>
<td>0%</td>
<td>100%</td>
</tr>
<tr>
<td>8chan</td>
<td>0%</td>
<td>0</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Voat</td>
<td>0%</td>
<td>0</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Overall</td>
<td>100%</td>
<td>176</td>
<td>41%</td>
<td>73%</td>
</tr>
</tbody>
</table>

Trolling volume was next examined to determine the degree of trolling during the past six months. Table 4 depicts that with respect to those utilizing the given social media site, the largest troll quantity per user includes YikYak (2.0 per user), Twitter (1.7 per user), Instagram (1.1 per user), Facebook (.8 per user), and Snapchat (.7 per user). YouTube and Reddit users were trolled .4 times per person, Google+ and other were trolled .1 times per person, and all other users were trolled 0 times. With respect to others being trolled on the given social media site, the largest troll quantity observed per user includes Twitter (131.5 per respondent), Facebook (61.8 per respondent), and Instagram (59.8 per respondent).
TABLE 4
Trolling Volume Per Student During Last 6 Months

<table>
<thead>
<tr>
<th>Social Media Site</th>
<th>Quantity of Student Trolls For Those Utilizing Site</th>
<th>Quantity of Other Person Trolls</th>
<th>Quantity For Only Students That Have Been Trolled</th>
</tr>
</thead>
<tbody>
<tr>
<td>Snapchat</td>
<td>0.7</td>
<td>6.1</td>
<td>5.8</td>
</tr>
<tr>
<td>Instagram</td>
<td>1.1</td>
<td>59.8</td>
<td>5.8</td>
</tr>
<tr>
<td>Facebook</td>
<td>0.8</td>
<td>61.8</td>
<td>5.7</td>
</tr>
<tr>
<td>Twitter</td>
<td>1.7</td>
<td>131.5</td>
<td>5.3</td>
</tr>
<tr>
<td>YouTube</td>
<td>0.4</td>
<td>18.0</td>
<td>5.4</td>
</tr>
<tr>
<td>LinkedIn</td>
<td>0</td>
<td>0.1</td>
<td>1.0</td>
</tr>
<tr>
<td>Pinterest</td>
<td>0</td>
<td>0.1</td>
<td>0</td>
</tr>
<tr>
<td>Google+</td>
<td>0.1</td>
<td>0.6</td>
<td>2.0</td>
</tr>
<tr>
<td>Reddit</td>
<td>0.4</td>
<td>11.2</td>
<td>5.0</td>
</tr>
<tr>
<td>Tumblr</td>
<td>0</td>
<td>7.1</td>
<td>0</td>
</tr>
<tr>
<td>YikYak</td>
<td>2.0</td>
<td>10.3</td>
<td>3.2</td>
</tr>
<tr>
<td>Other</td>
<td>0.1</td>
<td>1.4</td>
<td>1.0</td>
</tr>
<tr>
<td>4chan</td>
<td>0</td>
<td>11.5</td>
<td>0</td>
</tr>
<tr>
<td>8chan</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Voat</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Overall</td>
<td>8.2</td>
<td>222.9</td>
<td>19.8</td>
</tr>
</tbody>
</table>

The lesser trolled sites include YouTube (18 per respondent), 4chan (11.5 per respondent), Reddit (11.2 per respondent), and YikYak (10.3 per respondent). The least trolled sites include Tumblr (7.1 per respondent), Snapchat (6.1 per respondent), other (1.4 per respondent), Google+ (.6 per respondent), LinkedIn (.1 per respondent), and Pinterest (.1 per respondent). Finally, the quantity of trolls per student trolled was calculated. For those trolled, the highest volume sites include Snapchat (5.8 per victim), Instagram (5.8 per victim), Facebook (5.7 per victim), YouTube (5.4 per victim), Twitter (5.3 per victim), and Reddit (5.0 per victim). The lesser volume sites include YikYak (3.2 per victim), Google+ (2.0 per victim), LinkedIn (1.0 per victim), and other (1.0 per victim). All remaining sites had zero incidences. Overall, there were 8.2 trolls per student for users, 222.9 observed trolls per user, and 19.8 trolls for victims of social media sites during the past six months.

Finally, Spearman Rho correlations were calculated to determine if there are correlations between user minutes on a site and the quantity of student trolls or the quantity of other trolls per user. Given that there are five social media sites utilized by most students, only these sites were examined. Table 5 demonstrates that only one site, Snapchat, had a statistically significant correlation (significant at the .05 level) between minutes and the quantity of trolls per student. There were no significant correlations for any other sites. With respect to the quantity of other individual’s trolls, Twitter and YouTube had statistically significant correlations.
Case and King

(significant at the .05 level) between minutes and the quantity of trolls. There were no significant correlations for any other sites.

**TABLE 5**

Spearman Rho Correlations Between Minutes and Troll Volume Per Student by Social Media Site

<table>
<thead>
<tr>
<th>Social Media Site</th>
<th>Students Trolled</th>
<th>Others Trolled</th>
</tr>
</thead>
<tbody>
<tr>
<td>Snapchat</td>
<td>-.405*</td>
<td>.163</td>
</tr>
<tr>
<td>Instagram</td>
<td>-.036</td>
<td>.086</td>
</tr>
<tr>
<td>Facebook</td>
<td>-.101</td>
<td>.118</td>
</tr>
<tr>
<td>Twitter</td>
<td>.079</td>
<td>.185*</td>
</tr>
<tr>
<td>YouTube</td>
<td>.082</td>
<td>.353*</td>
</tr>
</tbody>
</table>

* Correlation is significant at .05 level (2-tailed).
** Correlation is significant at .01 level (2-tailed).

**CONCLUSIONS AND FUTURE RESEARCH**

Results demonstrate that undergraduates are active in social media with 100% of students visiting at least one site daily. Although there are nearly 20 sites identified in the study, most undergraduates visit five primary sites. These include Snapchat (95% of students), Instagram (88% of students), Facebook (81% of students), Twitter (76% of students), and YouTube (67% of students). Not only are these sites those with the largest student daily visits, they are also the sites that are visited for the most minutes each day. For example, undergraduates estimate spending 44 minutes per day on Twitter, 43 minutes per day on Instagram, 42 minutes per day on Snapchat, 39 minutes per day on YouTube, 32 minutes per day on Facebook, 24 minutes per day on Google+, 17 minutes per day on Tumblr, and 10 minutes per day on YikYak.

The ranking of the five primary social media sites used by undergraduates, however, are different from those of schools and businesses. The AACSB study found that Schools of Business use Facebook (65% of schools), Twitter (57% of schools), YouTube (42% of schools), and LinkedIn (32% of schools). Moreover, the Fortune 500 study found businesses use LinkedIn (97% of the firms), Twitter (76% of firms), Facebook (74% of firms), and YouTube (67% of firms).

When examining trolling, it is apparent that trolling occurs in nearly all social media sites visited by students and that undergraduates have been both victims and witnesses to others being trolled. For instance, 32% of respondents indicated being trolled at least once during the past six months on Twitter. Moreover, 19% were trolled on Instagram, 14% were trolled on Facebook, and 13% were trolled on Snapchat. In addition, 67% reported seeing others being trolled on Twitter and 64% noticed others being trolled on Facebook. Overall, 41% of students indicate being trolled and 73% indicate seeing another person being trolled during the past six
months. This is nearly identical to the Pew Research Center survey that found that 40% of adult Internet users were personally harassed online and 73% that had seen someone harassed online (Duggan, 2014).

In terms of volume, on average, a student is trolled more than one time per month (8.2 times per student during the past six months). In addition, each student reported seeing an average of 37 trolls per month (222.9 during the past six months) of others being trolled. When examining the quantity of trolls per student for only those trolled at the given web site, the average rises to more than three trolling incidences per month per user. The largest number of trolls per person during the six month period includes Snapchat (5.8 trolls), Instagram (5.8 trolls), Facebook (5.7 trolls), YouTube (5.4 trolls), and Twitter (5.3 trolls). The largest number of trolls per person of others includes Twitter (131.5 trolls), Facebook (61.8 trolls), and Instagram (59.8 trolls).

Finally, there are mixed results with regard to the relationship between online social media usage time and trolling. Only Snapchat had a significant correlation (at the .05 level) between site minutes and the volume of trolls per student. There were no statistically significant correlations with regard to Instagram, Facebook, Twitter, and YouTube. In terms of others being trolled, the only significant correlations (at the .05 level) between site minutes and the volume of trolls were with regard to Twitter and YouTube.

There are three important implications from the study. One finding is that trolling is a common phenomena both with respect to students and social media users in general. This suggests that measures may need to be implemented to help victims adequately deal with the pressures and negative issues such as depression that may be associated with being trolled. In addition, proactive measures could be used to educate students so they do not become a future troll themselves either because of retaliation or other desires.

A second implication is a result of the finding that trolling is common across nearly all of the social media sites identified in the study. Given both the percentage of victims and volume of trolls, it is apparent that the level of control by the site owners with regard to trolling is not adequate. This suggests that the social media sites need to be more vigilant in reviewing posts and user activity and likely more aggressive in security efforts to block or ban those that troll.

A third implication relates to the transitioning of undergraduate student social media usage from the traditional mainstay, Facebook, to newer platforms such as Snapchat and Instagram. It appears that Facebook is becoming passé for college students as evidenced in the both the percentage of active users and utilization time per day. Snapchat is used by 17% more users and Instagram by 9% more users than Facebook. Moreover, students use both Snapchat and Instagram by more than 32% more minutes each per day than Facebook. These findings imply that businesses
and universities may need to refocus their marketing and communication channels if they wish to effectively target the population of individuals that are younger than millennials. In addition, given this shift, there may be first mover advantages for these organizations.

The limitations of this study are primarily a function of the sample, sample distribution, and type of research. The use of additional universities and more equal distribution among academic class and gender would increase the robustness of results. Another limitation relates to the self-reported nature of the survey. Future research is needed to explore if gender and academic class affect online behavior and trolling incidence and to explore which measures may be implemented in the education process to bring about a positive change in behavior and response to victimization.

REFERENCES


A PATH FROM JOB AUTONOMY TO ORGANIZATIONAL CITIZENSHIP BEHAVIOR: THE ROLE OF PERCEIVED ORGANIZATIONAL POLITICS AS MEDIATOR

Tae Seok Yang  
Western Illinois University  
Alankrita Pandey  
Eastern Michigan University  
Yin-Chi Liao  
Joseph J. Dobson  
Western Illinois University

ABSTRACT: Job autonomy and organizational citizenship behavior have both been the focus of research in the organizational literature, yet few attempts have been made to investigate the relationship between the two. This study aimed to further the body of knowledge by applying a theory-driven approach to examine how job autonomy affects organizational citizenship behavior. Building on the research surrounding job autonomy and organizational citizenship behavior, as well as social exchange theory, the authors proposed a model in which perception of organizational politics and organizational commitment mediate the impact of job autonomy on two forms of organizational citizenship behavior—altruism, and conscientiousness. The mechanism we used examined these impacts in series. A conditional process analysis with a bootstrapping technique suggested a partial mediation model. The results showed that job autonomy significantly reduced employee’s perceptions of organizational politics, and this decreased perception improved employees’ engagement in commitment toward their organizations and increased conscientiousness. It is of interest to note that this concurs with Hackman’s model regarding work outcomes. While the results showed effects on conscientiousness, there were not similar effects on altruism. The mixed nature of our results indicates there are likely other moderating factors which need to be considered in future models, such as personality and competitive climate.

Key words: Job Autonomy, OCB, Perceived Organizational Politics, Mediation

INTRODUCTION: Organizational citizenship behavior (OCB) has been receiving a lot of attention lately. While task performance is directly related to work outcomes, citizenship behavior has been found to provide important work outcomes by enhancing the social and psychological work environment (Mackenzie, Podsakoff, & Fetter, 1993; Organ, 1997). Because of its positive contributions to the
organization, managers often take into account citizenship when conducting employee performance evaluations (Podsakoff & Mackenzie, 1994). George and Jones (1997) suggested that organizational context such as skill level, role definition and interpersonal relationships plays an important role in determining whether employee engage in behaviors that go over and beyond their prescribed job roles.

The current study extends this examination of work contexts by examining job autonomy and its relationship with citizenship behavior. Previous job autonomy research has found a small direct relationship of autonomy with performance (Fried & Ferris, 1987; Liden, Wayne, & Sparrowe, 2000). Similarly, studies that investigated the relationship between job autonomy and OCB have reported mixed results. Some researchers found a significant direct relationship of job autonomy with citizenship behavior (Farh, Podsakoff, & Organ, 1990; Capelli & Rogovsky, 1998; Chen & Chiu, 2009; Krishnan, Ismail, Samuel, & Kanchymalay, 2013) while other researchers found a non-significant direct relationship between two variables (Chiu & Chen, 2005; Schaubroeck & Fink, 1998; Todd & Kent, 2006). These findings suggest that there exist possible mediators that cloud the direct relationship between job autonomy and OCB, and that exploring possible mediators would enhance our understanding of their relationship. The current study proposes a model in which organizational politics and organizational commitment mediate the relationship between the two.

JOB AUTONOMY TO CITIZENSHIP BEHAVIOR THROUGH ORGANIZATIONAL POLITICS: Organizational citizenship behavior (OCB) is defined as employees’ extra-role behaviors beyond what are formally written in job description. A meta-analysis on organizational citizenship behavior (Podsakoff, MacKenzie, Paine, & Bachrach, 2000) suggests that task variables have an important impact on OCB. Following that assessment, job tasks that motivate employees and create in their minds a sense of meaning would therefore serve as antecedents to OCB. Hackman and Oldham (1975) introduced job autonomy as a job/work contextual factor along with the other factors of job enrichment. Job autonomy is defined as ‘the degree to which the job provides substantial freedom, independence, and discretion to the individual in scheduling the work and in determining the procedures to be used in carrying it out’ (Hackman & Oldham, 1975, p. 258).

According to Organ (1988), OCB is individual discretionary behavior that is not explicitly required by the organization, and thus, individuals perform OCB as a result of personal choice. On the other hand, task performance is formally recognized and required by the formal system. Therefore, the individuals normally have greater discretion in exerting OCB than task behaviors (Organ, 1988). Also, individuals who are given more discretion in their jobs will be more likely to engage in citizenship behaviors (Morgeson, Delaney-Lkinger, & Hemingway, 2005). Increased autonomy provides greater discretion to employees in how to perform their work and allows employees to have a broadened definition of role (Troyer, Mueller, & Osinsky, 2000), and employees with increased discretion further increases their engagement in OCB (Morgeson et al., 2005). Similarly, Blau’s
(1964) social exchange theory also forms the basis for the relationship between job autonomy and OCB. For instance, jobs that are autonomous provide employees with the opportunity to make decisions about their work and in return they reciprocate with positive work behaviors such as OCB directed towards the organization (Organ, 1997).

Some researchers found a significant direct relationship of job autonomy with citizenship behavior (Farh et al., 1990; Capelli & Rogovsky, 1998; Chen & Chiu, 2009; Krishnan et al., 2013). However, some studies also reported non-significant direct relationships between the two (e.g., Chiu & Chen, 2005; Schaubroeck & Fink, 1998; Todd & Kent, 2006). Typically, job autonomy studies acknowledged a small direct relationship of autonomy with performance (Fried & Ferris, 1987; Liden et al., 2000). This suggests that there exist possible mediators that cloud the direct relationship between job autonomy and OCB, and that exploring possible mediators would enhance our understanding of their relationship. The current study proposes a model in that another contextual variable, organizational politics, mediates the relationship between the two.

**JOB AUTONOMY TO PERCEIVED ORGANIZATIONAL POLITICS:** Organizational politics is defined as the process of using influence attempts that are not required by a formal role in the organization to maximize self-interest at the expense of the organizational goals (Ferris & Kacmar, 1992; Ferris, Russ, & Fandt, 1989). Theoretically it is believed to hinder the employee-organization exchange relationship (Aryee, Chen, & Budhwar, 2004). For instance, in a highly political environment, rewards and promotions are decided by subjective factors such as personal relationship and power so that employees believe their performance is weakly tied to desired outcomes (e.g., rewards). Therefore, organizational politics creates a detrimental impact on the organizational processes (e.g., decision making, promotion) and eventually on the organizational outcomes (Vigoda, 2000).

Employees engage in more political behaviors as uncertainty or ambiguity in the work place increases (Ferris et al., 1989; Fandt & Ferris, 1990). Several empirical studies found that dysfunctional behaviors in organizations occurred due to ambiguous goals and unclear performance measurement (Bevan & Hood, 2006; Speklé & Verbeeten, 2014). Therefore, the most efficient way to reduce political behaviors in the organization may be to reduce role ambiguity and unclear performance evaluation procedures. Daft (1989) argued that lack of autonomy creates feelings of powerlessness among employees so they perceive more organizational politics. In addition, Ferris and Kacmar (1992) suggested that job/work environment, such as job autonomy, reduce perceptions of politics among employees because it reduces uncertainty in the work environment. A sociotechnical systems perspective (Pasmore, 1982) provides further theoretical explanation for why organizational politics is reduced by job autonomy. It argues that by increasing job autonomy, organizations reduce employees’ dependence on managerial monitors and control, which lessens bureaucratic constraints imposed by managers. Under the greater managerial constraints, employees may think their
evaluations are determined based on personal relationships. As a result, they are likely to perceive the workplace as political and engage in political behaviors in order to secure their self-interest. When greater autonomy is given, employees will spend less time and mental energy worrying about what managers might do to them and will see no need to exert political behaviors. Eventually, they will less likely perceive the organization as political.

In our study, we used perceptions of organizational politics instead of using an objective measure of organizational political behaviors following suggestions of previous studies (e.g., Ferris & Kacmar, 1992; Ferris et al., 1989; Gandz & Murray, 1980). The rationale behind this approach is that people do not respond to the actual reality, but respond based on perceptions of reality (Lewin, 1936). For this reason, in research of work environments, contextual variables have been typically defined on the basis of perceived attributes (Ferris & Kacmar, 1992). Therefore, it is appropriate to understand organizational politics as a subjective experience and state of mind (Gandz & Murray, 1980) and to use perceptions when referring to organizational politics as a contextual variable for study (Ferris et al., 1989). Therefore, we suggest the first hypothesis as following:

**Hypothesis 1:** Job Autonomy will be negatively related to Perceptions of Organizational Politics.

**PERCEIVED ORGANIZATIONAL POLITICS TO ORGANIZATIONAL COMMITMENT:** Chang, Rosen, and Levy (2009) classified organizational outcomes into two categories: immediate outcomes (attitudes) and distal outcomes (task performance, OCB and turnover intention). They also proposed a mediation model where perceived organizational politics indirectly impacts distal outcomes through immediate outcomes. Cropanzano, Howes, Grandey, and Toth (1997) provided a support for this linkage with empirical evidence that shows employees evaluated the employee-organization exchange relationship more negatively as strain increases. Similarly, Droy (1990) suggested that employees’ negative attitude to their perception of organizational politics is more immediate and it indicates their negative behavioral responses in the long run.

Based on the work stress framework of Schaubroeck, Cotton, and Jennings (1989), Chang et al. (2009) conceptualized perceptions of organizational politics as a hindrance stressor which impedes individual’s work-related achievements. They argued that since these constraints are coming from a high degree of perceived organizational politics, employees’ ability to achieve their career goals is impeded. As such, they experience psychological strain in the work environment and consequently build up negative attitudes toward their organization and job. In this study, we used organizational commitment as the attitudinal work outcome because this variable is theoretically considered as a fundamental construct of job attitudes (Mowday, Steers, & Porter, 1979). Organizational commitment is the individual’s psychological attachment to his or her organization by identifying with the values of the organization and being deeply involved in the organizational activities (Mowday et al., 1979). If employees see themselves unfairly treated by the organization, they are likely to reduce their attachment to the organization.
Researchers consistently found the negative relationship between organizational politics and affective commitment (e.g., Kimura, 2013; Maslyn & Fedor, 1998; Nye & Wit, 1993; Vigoda, 2000).

Hypothesis 2: Perception of Organizational politics will be negatively associated with organizational commitment.

ORGANIZATIONAL COMMITMENT TO OCB: Weiner (1982) conceptualized organizational commitment as the internalized belief to act for organizational interests. Therefore, highly committed individuals feel responsible for personal sacrifice for sake of the organization instead of deciding their behaviors primarily based on exchange relationships (i.e., rewards, punishment). Thus, employees with strong affective commitment are more likely to behave on behalf of the organization and devote more effort at work (Meyer, Stanley, Herscovitch, & Topolyntsky, 2002). Since OCB represents discretionary behaviors that occur even though they are not related to the formal reward system, it is reasonable to anticipate a positive linkage between commitment and OCB. There have been a plethora of empirical findings to support commitment as an antecedent of OCB (e.g., O’Reilly & Chatman, 1986; Williams & Anderson, 1991). In addition, the meta-analyses conducted by Organ and Ryan (1995) and Meyer et al. (2002) showed the same results and supported the positive relationship between affective commitment and OCB. Therefore, we propose the following hypothesis.

Hypothesis 3: Organizational commitment will be positively related to OCB.

Following the above discussions from hypothesis 1 through hypothesis 3, we suggest a two-stage mediation model where job autonomy impacts OCBs indirectly through organizational politics (i.e., contextual variable) and organizational commitment (attitudinal variable) in series.

Hypothesis 4: Job autonomy impacts OCB indirectly through perception of organizational politics and organizational commitment in series.

METHOD: Data was collected through an online survey from 121 respondents who were currently employed. However, 14 cases were removed due to missing responses, resulting in 107 valid cases. 41% of the sample were male. 65% were white followed by black (18%) and Hispanic (16%). Approximately 83% of participants were permanent employees (either full-time or part-time). The average tenure was 7.85 years (SD=9.38) and participants worked 42.5 hours a week on average (SD = 9.11).

The major constructs are measured by the scales in the existing literature. All measures except job autonomy consisted of a 5-point Likert-type scale from strongly disagree (1) to strongly agree (5). Job Autonomy: Job autonomy was measured with a five-item scale developed by Sims, Szilagyi, and Keller (1976). The sample items include “How much are you left on your own to do your own work?” and “To what extent are you able to act independently of your supervisor in performing your job function”. Participants were asked to rate each item based on a five-point Likert scale (1=very little to 5=very much). Cronbach alpha was .78. Organizational politics: Organizational politics was measured by a scale of
perceived organizational politics developed by Maslyn and Fedor (1998). The items intend to ask the participants the degree to which they perceive their work environment as politics. By examining Cronbach’s alpha if item deleted, two items that caused internal inconsistency were removed. Cronbach alpha of the remaining items was .76. **Organizational commitment**: We measured organizational commitment by a six-item scale developed by Marsden, Kalleberg and Cook (1993). A sample item of this scale is “I would take almost any job to keep working for this organization.” Cronbach alpha was .71. **Organizational citizenship behavior (OCB)**: Since OCB is multidimensional and different types of OCB affect organizational effectiveness in different ways, it is necessary to distinguish types of OCB for better understanding. Therefore, we measured two aspects of OCB, altruism and general compliance (i.e., conscientiousness), following Smith, Organ, and Near (1983). Altruism is defined as discretionary behavior such as helping colleagues with tasks or problems directly or indirectly related to organizational performance. General compliance is another form of discretionary behavior that helps organizational efficiency such as less absenteeism or taking less time for breaks. Altruism was measured with a five-item scale and general compliance was measured with five items from the conscientiousness scale developed by Podsakoff, Mackenzie, Moorman, & Fetter (1990). Cronbach alphas were .72 and .70, respectively.

<p>| TABLE 1. MODEL COEFFICIENTS FOR THE HYPOTHESIZED TWO-STAGE MEDIATION MODEL ANALYSIS |</p>
<table>
<thead>
<tr>
<th>Consequent</th>
<th>M1</th>
<th>M2</th>
<th>Y (OCBA)</th>
<th>Y (OCBC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IVs</td>
<td>Coeff. (SE)</td>
<td>Coeff. (SE)</td>
<td>Coeff. (SE)</td>
<td>Coeff. (SE)</td>
</tr>
<tr>
<td>Constant</td>
<td>3.63 (.39)**</td>
<td>3.24 (.48)**</td>
<td>3.31 (.52)**</td>
<td>.2.69 (.48)**</td>
</tr>
<tr>
<td>JA</td>
<td>-.20 (.10)**</td>
<td>.32 (.09)**</td>
<td>.13 (.09)</td>
<td>.19 (.08)*</td>
</tr>
<tr>
<td>M1</td>
<td>-.36 (.09)**</td>
<td>-.03 (.09)</td>
<td>.08 (.08)</td>
<td></td>
</tr>
<tr>
<td>M2</td>
<td></td>
<td>.12 (.09)</td>
<td></td>
<td>.16 (.08)*</td>
</tr>
<tr>
<td>R² = .04</td>
<td>R² = .51</td>
<td>R² = .07</td>
<td>R² = .51</td>
<td></td>
</tr>
<tr>
<td>F (1, 105) = 4.10</td>
<td>F (2, 104) = 17.84</td>
<td>F (3, 103) = 2.76</td>
<td>F (2, 103) = 17.84</td>
<td></td>
</tr>
<tr>
<td>(P &lt; .05)</td>
<td>(P &lt; .01)</td>
<td>(P &lt; .05)</td>
<td>(P &lt; .01)</td>
<td></td>
</tr>
</tbody>
</table>

*Note: JA=Job autonomy; M1 (mediator 1) = POPS (Perceived Organizational Politics); M2 = Commitment; OCBA = Altruism; OCBC = Conscientiousness; * P < .05 ** P < .01*

**ANALYSIS AND RESULTS:** We applied conditional process analysis by using the PROCESS macro for SPSS (Hayes, 2013) following the recommendations from Preacher, Rucker, and Hayes (2007). This statistical analysis approach enables us to explore every possible path to test direct and indirect effects simultaneously and more importantly, to examine the integrated analytical model instead of examining a particular relationship between two variables at a time. We conducted 1000 resampling bootstraps to estimate effect sizes of each path and constructed 95% bias-corrected confidence intervals for each statistic estimate.
As represented in Table 1, job autonomy is negatively related to employees’ perception of organizational politics ($\beta = -0.20, p < .01$) and perceived organizational politics has a negative relationship with commitment ($\beta = -0.36, p < .01$), supporting hypotheses 1 and 2. However, hypothesis 3 received partial support because commitment is significantly positively related to conscientiousness ($\beta = 0.16, p < .05$) but its relationship with altruism is insignificant ($\beta = 0.12, p > .05$). We also hypothesized significant relationships between variables through the full mediation path (Hypothesis 4).

Table 2 shows the summary of the total, direct, and indirect effects of job autonomy on OCBs through organizational politics and commitment in series. Overall, it was found that organizational politics and commitment do not mediate the relationship between job autonomy and altruism, but they mediate the path from job autonomy to conscientiousness (mediation effect size = 0.01, 95% CI = 0.00 to 0.04). The 95% confidence intervals for all of the indirect effects of job autonomy to altruism include zero, which indicates that there are no significant mediation effects of organizational politics and commitment on the path from job autonomy to
altruism. Also, job autonomy did not directly affect altruism (95% CI = -.04 to .30). On the other hand, job autonomy has not only a direct effect on conscientiousness (95% CI = .03 to .34) but also indirectly influences conscientiousness either through commitment only, without going through organizational politics (indirect model 5: 95% CI = .01 to .14) or through both organizational politics and commitment in series (indirect model 6: 95% CI = .0001 to .0442). Interestingly, as shown in the indirect models 4 and 6, organizational politics does not mediate the relationship between job autonomy and conscientiousness alone (i.e., its confidence interval includes zero) but only with commitment in series. As additional evidence shows in table 1, job autonomy affects commitment ($\beta = .32$, $p < .01$) and conscientiousness ($\beta = .19$, $p < .05$), indicating a partial mediation model.

**FIGURE 1.** Standardized path coefficients for the indirect effects of Job autonomy (JA) on altruism (OCBA) and conscientiousness (OCBC) through perceived organizational politics (POPS) and organizational commitment (OC) in series. * $p < .05$; ** $p < .01$

**DISCUSSION:** The current study attempted to examine a mechanism which examines the indirect impact of job autonomy on organizational citizenship behavior through organizational politics and organizational commitment in series (see Figure 1). Specifically, we tried to examine how job autonomy is related to particular kinds of OCBs such as altruism and conscientiousness. Our findings partially supported our full two-stage mediation model (Hypothesis 4). Contrary to our belief, job autonomy was not an immediate predictor of altruism nor indirectly affected it via perceived organizational politics and commitment. On the other hand,
job autonomy affected conscientiousness indirectly via two attitudinal variables in series, which supported our two-staged mediation model. Therefore, Hypothesis 4 was supported only in the job autonomy – conscientiousness link. More specifically, job autonomy significantly reduced employees’ perception of organizational politics (supporting Hypothesis 1) and low perceived organizational politics enhanced employees’ commitment to the organization (supporting Hypothesis 2). Increased commitment encouraged employees to engage in conscientiousness but not altruism (partially supporting hypothesis 3).

These findings are similar to some previous studies and different from others. For instance, Farh et al. (1990) found the positive relationship between job autonomy and altruism but they did not find a significant relationship of job autonomy with conscientiousness. On the other hand, Bell and Menguc (2002) reported that job autonomy was related to both altruism and conscientiousness. Therefore, future studies should continue to investigate the relationships between job autonomy and various OCB dimensions by incorporating possible mediators and moderators.

From Hackman’s model, we know that increased levels of the core job dimensions (of which autonomy is a component) are associated with less absenteeism and employee turnover. The present study shows the link between job autonomy and employee’s commitment to the organization. An increase in commitment to the organization should result in better attendance by employees, as well as longer tenures of employment. Thus we see conjoint support of outcomes between Hackman’s model and the model presented here. Moreover, our study demonstrate that increasing job autonomy leads to other beneficial outcomes to the organization, which are decreased perceived organizational politics, increased commitment, and increased OCB. These findings suggest that while there may be risks with increasing job autonomy, if the downsides are mainly reduced performance during a relatively short learning period and egos of supervisory personnel, then the risks are well worth the change in the status quo.

There are several limitations of the current study. First, data was collected from a small sample (N=107) and all variables were measured by self-report survey. Thus, the relationships between variables might have been inflated due to response bias caused by a small sample size and self-reported measurement. We examined a possible significant common method effect by running Harman’s single factor test as recommended by Podsakoff and Organ (1986). This analysis result showed no significant effect of common method bias on our findings. However, further research should address this issue by incorporating multiple sources of data. Second, we applied a concurrent data collection design. In this approach, it is more likely to find higher correlations between variables and does not fully explain a longer-term effect of the predictor. Therefore, future researchers may consider employing a longitudinal design. Third, we looked at only one aspect of performance, OCB. In addition, future researchers are recommended to include task performance and dysfunctional behavior along with OCB in their models (Rotundo & Sackett, 2002).
REFERENCES


ABSTRACT: In professional sports teams that win games consistently do well financially. Over the years, profits have motivated professional sports teams to search for good players. In their pursuit of wins, teams have recruited talented players regardless of their race or national origin. This paper examines the effects of hiring a diverse set of players on team success in the National Basketball Association (NBA). The effects of employee heterogeneity on team performance are tested using panel data from the 2002-2008 NBA seasons. This paper shows the benefits and problems associated with employing a geographically diverse workforce in the NBA. The Herfindahl-Hirschman Index (HHI), a measure of market power, is adapted to measure team diversity. Win percentages of teams are regressed against their HHIs and other control variables such as player talent and coaching expertise. The results show that increases in diversity as measured by the HHI and international player’s minutes lead to higher win percentages for their teams.

Key Words: Heterogeneity, National Basketball Association, human resource management.

1. Introduction

Since competition among companies is fierce, organizations continually seek to maximize performance given capital constraints (Giambatista, 2004; Murray 2010). The human capital constraint is of specific interest given its changed nature in recent years, as companies have shifted organizational forms and structure. In particular, bureaucratic organizations with strict lines of communication and a hierarchical leadership structure, are not always suited for the contemporary challenges that they face (Storey, 2009; Cunningham, 2007). Thus, one should be wary of dramatically overstating organizational change given the nature of labor markets (Storey, 2009). With the increase in the diversity and breadth of talent pools globally, as well as changes in technology, the competitive landscape for companies is as fierce as it has ever been (Fink & Pastore, 1999; Erhardt, 2003; Lazear 1999).
This emerging landscape makes it critical to understand, how to employ a diverse or heterogeneous workforce. Potential international employees are competing for, and winning, highly skilled jobs that were previously inaccessible to them (Storey, 2009; Hankin, 2005; Lazear, 1999). Global firms are the ones that build workforces from employees coming from diverse sets of cultures and countries (Lazear, 1999). These firms are seeing an influx in high-level international employees in the technology, pharmaceutical and healthcare industries (Storey, 2009). There has also been an influx of international talent in the professional sports industry (Fenn, 2005). Human resource management teams (HR) in these industries understand that recruiting and training a new employee comes with certain sunk costs. These sunk costs, of time and money, often increase when HR has to cater to the legal, lingual and cultural differences of employees from diverse backgrounds (Lazear, 1999; Cunningham 2007). Given the competitive nature of most of these industries, a relevant question is whether the added benefits from a culturally heterogeneous workforce justify the startup and acquisition costs associated with these workers? (Lazear, 1999; Simmons & Berri, 2011; Surdam, 2012)

1.1 Striving for a balance. Organizations that effectively address the question of whether to employ heterogeneous workers or not, may experience a competitive advantage in their field (Cunningham, 2007; Lazear, 1999). In order to be successful, HR needs to convey the positive and negatives which can stem from a diverse international group. Cultural heterogeneity among workers has been shown to contribute two valuable assets, diversity in thought and a unique skill set (Lazear, 1999). This can lead to increased creativity and innovation in problem solving, as well as fostering a more open work environment (Cunningham, 2007; Lazear, 1999).

In contrast, too much heterogeneity in the workforce has the potential to lead to communication deficiencies or to resentment among coworkers (Kahane et al., 2013). Today, increasing diversity is often suggested as a universal antidote to many corporate problems. However, understanding the dichotomy of diversity makes selecting the proper diversity management technique essential for global firms (Lazear, 1999; Cunningham, 2007; Storey, 2009; Kahane et al., 2013). The lack of quantitative measurements or statistics on diversity are challenging for global firms (Kahn, 2000; Lazear, 1999; Ottaviano, 2006). This stems from the limited availability of performance and demographic statistics for most industries, (Kahane et al., 2013; Kahn, 2000; Lazear, 1999; Ottaviano, 2006). In most industries measuring firm output as a function of worker diversity is extremely challenging, if not impossible, (Kahane et al., 2013; Kahn, 2000; Lazear, 1999).

2. Literature Review
Unlike most professional industries, such as law or medicine, the professional sports industry in North America is an exception to acquiring worker and firm data. It is an industry where, as Kahn (2000) notes, “…we know the name, face and life history of every production worker and supervisor in the industry.
Total compensation packages and performance statistics for each individual are widely available.” Furthermore, the professional sports sector in North America has seen a dramatic increase in foreign employees over the past two decades (Eschker, 2004). The sports industry can answer empirical labor market questions that other industries cannot because of the availability of data on employee diversity and their performance (Kahn, 2000).

Scully’s seminal work estimated a production function of wins in sports (Scully, 1974). Scully models the winning percentage of a Major League Baseball (MLB) team, as a function of worker, managerial and team skills. Kahane built upon this general model by adding team diversity to examine the potential gains of employing culturally diverse work teams. He used panel data from the National Hockey League (Kahane, 2013). Kahane’s study examined the specific distribution of players from geographic regions using a Herfindahl-Hirschman index (HHI). Measures of this index typically lie between zero and one. A value of zero indicates that the work force is fairly diverse and that roughly equal numbers of players come from each ethnic group. However, a value closer to one indicates that most of the workers come from a few ethnic groups.

2.1 Factors Affecting Output. The existing literature has examined the effects of other factors on firm output (wins) in the NBA such as change in coaching status and pay inequality (Hofler & Payne, 2006; Martinez, 2013; Simmons & Berri, 2011). Kahn highlighted the negative effects of discrimination, based on race, in the NBA during the 1980s, (Kahn 1988). In the 1990s, international players were paid a premium over equivalent domestic players for the 1996-1997 and 1997-1998 seasons. (Eschker, 2004). A dramatic increase in international players during this time led to what Eschker called the ‘winner’s curse.’ Teams overpaid for players before they had developed resources to accurately scout and evaluate them (Jozsa, 2011; Eschker, 2004). This example shows the importance of why having a solid HR management plan relating to internationally diverse players is important in any industry.

Previous research using NBA data from 1950-1997 has revealed that an increase in race and age diversity in the NBA had a negative impact on teams’ winning percentages. However, these two variables did not affect winning percentages in Major League Baseball (MLB) over the same time period, (Timmerman, 2000). The difference in these results may be due to the fact that basketball is a complex invasion sport, which requires a relatively higher degree of teamwork and coordination between players than baseball. A complex invasion sport is one where players must work together to advance an object (a ball or puck) into the other team’s territory or goal. Baseball is more a sum of individual performances. Whether or not a MLB player gets a hit rarely, if ever, depends upon his teammates direct actions (Timmerman, 2000). In complex invasion sports, in order to produce wins, teammates must work together to score points and also to prevent opponents from scoring. Conflict and uncooperative behavior based on racial stereotypes seem more likely to affect performance in basketball, and other complex invasion sports. (Timmerman, 2000).
2.2 Benefits of Diversity. A recurrent theme throughout the literature is that the benefits associated with increased organizational diversity come with costs (Kahane et. al., 2013). Incorporating unique talents, perspectives and problem-solving methods into the workplace, which differ from domestic employees often require organizational adaptation to cultural and language differences, (Cunningham, 2007; Fink & Pastore, 1999). Since the mid-1990s, the National Basketball Association (NBA) has witnessed an explosion of recruited international talent (Eschker et al., 2004). During the 1999-2000 NBA season, there were only 35 “international “players, whereas in the 2012-2013 season, the number of international players increased to 83 from 36 different countries. International players accounted for 18% of players on all NBA rosters during the 2012-2013 season, (NBA, 2012). This trend of recruiting international talent has also permeated other highly competitive markets such as technology firms and pharmaceutical companies. In these industries talent searches are conducted at both domestic and international levels, (Storey, 2009).

We do recognize that extrapolating professional sports labor market data to other industries has its limits (Kahn, 2000). However, we argue that increasing our understanding of the relationship between increased diversity and organizational output can, indeed, be useful to any industry or firm looking to have a heterogeneous workforce (Hofler & Payne, 2006; Zak, Huang & Siegfried, 1979; Berri 2013).

2.3 Effects of Heterogeneity on Performance. This study builds on the theoretical framework proposed by Timmerman and Janis, (Timmerman, 2000; Janis, 1972). In keeping with Kahane, we use the Herfindahl-Hirschman Index (HHI) to measure the concentration of international players on NBA teams, (Kahane, 2013). This measure is typically a measure of market concentration but has been adapted by sports economists including Kahane to measure the distribution of international players among teams. One reason why the HHI is used by sports economists is that it is a continuous variable that measures the distribution of a given variable such as diversity. Thus it is easily employed in regression models. Next, we model regular season NBA win percentages as a function of diversity and other control variables using data from the 2002-2003 to 2008-2009 seasons.

3. Theory and Model

Scully highlighted that certain inputs in sports contribute more effectively to a team’s output. These inputs for sports, in general, are worker talent and managerial ability (Scully, 1974). In addition, we include team diversity as another input. The first two inputs serve as control variables. Therefore, the general function used to model a NBA team’s success in the regular season from 2002-2008 is given by equation (3.1).

\[
\text{Firm Performance}_{i,t} = f (\text{Worker Talent}_{i,t}, \text{Managerial Skill}_{i,t}, \text{Team Diversity}_{i,t})
\]

(3.1)

The subscript \(i\) refers to the year and the subscript \(t\) refers to the team. Firm performance is given by a team’s win percentage (WinPercentage\(_{i,t}\)) during the regular season, which is a common practice in the sports economics literature.
(Kahane et al., 2013; Timmerman, 2000; Giambatista, 2004; Fizel, 1997). Win percentage is given by the number of wins divided by the total number of games played during the regular season. During the years 2002-2008, all teams played in 82 games per season. A team wins a game in basketball by scoring more points than its opponent scores in the 48 minutes of playing time. In the event that the teams are tied at the end of 48 minutes, they play a five-minute overtime period. These overtime periods continue until one team has more points than the other team at the conclusion of overtime.

3.1 Worker talent. Worker talent is measured through a number of performance metrics recorded by the NBA. Individual player skill vectors were combined as weighted averages based on minutes played to produce team-by-team statistics. The team’s statistics were created for each of the following categories: Field goal percentage (FGP), free throw percentage (FTP), three point field goal percentage (TPFGP), rebounds per game (RPG), and turnovers per game (TOG). Each player's metric was weighted by the ratio of that player’s minutes to their team’s total minutes. Then these metrics were summed to give the overall team metric in an area such as FGP. The same process was repeated for all the other metrics mentioned above. The team, for the purposes of this study, was composed of players who played at least 984 minutes in a regular season, which averages out to 12 minutes a game. Players that changed teams were only included if they had been traded once and had played in at least half of the games during that season. In addition these players had to have played at least 984 minutes. Assuming players with very little playing time have less of an effect on win percentage, this was done to only include the significant players. Individual player skill vectors were combined as weighted averages based on minutes played to produce team-by-team statistics. We assume that players with more minutes have a greater effect on win percentage. These weighted team statistics were field goal percentage, free throw percentage, three-point field goal percentage, rebounds per game, and turnovers per game. We will conduct a test of means among player performance metrics by region. However, in the regression model of the production function we will summarize the level of worker talent by using a team’s relative salary ratio and its square as explained below.

3.2 Payroll. Payroll is the sum of all players’ salaries on the team. A player’s salary is assumed to be a measure (albeit imperfect) of individual talent. Talent is often measured by in-game statistics. However, talent, and a player’s salary, can also take into account softer, less quantifiable skills such as leadership ability or the effect a player has on the locker room (Kahane et al., 2013). It should be noted that payroll as a measure of team ability has its flaws, as individual salaries are negotiated in a restricted labor market (Kahane et al., 2013; Fenn, 2006). Furthermore, NBA team salaries are worth tens of millions of dollars, so even a relatively large numerical change in payroll will have a minor impact on win percentage.

For this reason, we use a salary ratio ($$\text{SALARYRATIO}_{i,t}$$), as an approximation of the talent level of a team. The salary ratio is given by team
payroll divided by the average league payroll. This ratio is used as an independent variable in the model. This variable attempts to measure the overall talent of a team in comparison to others. In keeping with Kahane, the squared value of the salary ratio \( \text{SALARYRATIOSQ}_{i,t} \), is also included as a regressor, Kahane, (2013). This squared salary ratio term allows for the possibility of diminishing returns of the talent level due to some fixed input such as the maximum amount of playing time. The team salary data is a sum of the entire rosters salaries, regardless of total minutes played. It is assumed, however, that the highest paid players are the ones playing the most minutes, so the salary data is a close proxy. It is hypothesized that, ceteris paribus, a higher value of \( \text{SALARYRATIO} \) will have a positive impact on win percentage, while a negative coefficient on the squared term, \( \text{SalaryRatioSQ} \), is indicative of concave down win production function. In other words, as the salary ratio increases with wins the linear term is expected to be positive whereas the squared term of salary ratio is expected to have negative coefficient indicating diminishing returns to the talent level (presumably due to some fixed input such as maximum playing time). Teams with nine excellent players and a higher than average aggregate payroll are often ineffective in engaging all talent equally to fully reap the benefits from the higher than average salaries because of the highly constrained playing time (opportunity for performance) with only 48 minutes in a regular game.

3.3 Coaching Skill. A coach is responsible for creating a game plan that will best organize and motivate his players and the team to win (Martinez, 2013). In this sense, a coach’s role is similar to a manager within an organization, who oversees and directs employees towards success. A coaches’ managerial skill is measured by calculating a coaches’ career win percentage, which follows aspects of previous studies measuring coaching ability in sports (Martinez, 2013; Kahane et al., 2013). Career win percentage \( \text{CAREERWINPCT}_{i,t} \), is calculated by dividing the total number of games won as coach by the total games coached. In cases where there has been a coaching change, we use the career win percentage of the coach with the higher number of total games coached in that season.

A basketball coach decides the amount of playing time a player gets, and what combination of players are on the court together. Since basketball has such a high degree of interdependence, it is important to put players on the court together who will succeed (Timmerman, 2000). This high degree of interdependence can lead to disputes and disagreements within a team, and it is often the job of a coach, as it would be for a manager in a firm, to settle them fairly and justly. A skilled coach will do these things more frequently, and thus have a better chance of winning games. Therefore, ceteris paribus, it is hypothesized that a higher career coaching win percentage will improve a team’s performance.

3.4 Heterogeneity. Diversity or heterogeneity models the international make-up of teams by employing the Herfindahl-Hirschman Index (HHI). The HHI is commonly used to measure the concentration of international players on a team. An international player was defined as a player who was born outside of the United States or Canada and did not play in the U.S. until at least age fifteen. For example, Loul Deng, was classified as an international player, even though he played one
year collegiately at Duke. Players are placed in one of six geographic regions. The regions are USA/Canada, Europe, Africa, South/Central America, Asia/Australia and Other. The other category was reserved for players not fitting any of the previous categories. Israel and Haiti are examples of two countries labeled as “Other.” This builds on the work of Kahane who used the HHI when measuring heterogeneity in the NHL. The formula for the HHI is given by equation (3.2).

\[ HHI = \sum (MS_{it}) \]  

(3.2)

MS_{it}, or market share, is the percentage of players on team \( t \) from regions \( i \) in a given year. Therefore, a team with no diversity, or all players from one region, would have an HHI=1 whereas a team with roughly equal proportions of players from all six regions will have a value closer to zero. The HHI is employed to record a more accurate depiction of diversity than past sports literature, where diversity was measured through dummy variables (Eschker & Perez, 2004). A more accurate depiction of diversity should be achieved by this regional approach rather than simply classifying players as either international or domestic. The HHI is calculated for every team for each of the years in the dataset.

Within the NBA, the majority of players are not international. For this reason, the HHI numbers will be skewed towards the domestic group. While the HHI accounts for the number of players from various regions, on a team, it does not account for their actual playing time. To account for this, international player’s minutes (INTLMINS_{it}), are also included instead of HHI in the model. This is in keeping with Kahane et al.’s fix to a similar problem in the NHL diversity metric, (Kahane et. Al., 2013).

Our simple model for modelling the regular season win percentage production function is given by equation (3.3).

\[
WINPCT_{it} = \beta_0 + \beta_1 \text{SALARYRATIO}_{it} + \beta_2 \text{SALARYRATIOSQ}_{it} + \beta_3 \text{CAREERWINPCT}_{it} + \beta_4 \text{HHI}_{it} + \beta_5 \text{INTLMINS}_{it} \epsilon_{it}
\]  

(3.3)

4. Analysis and Results

We employ both a test of means for player performance metrics by region and the regression model for the production function given by equation (3.3). Table 4.1 shows the summary statistics for our regression model. The panel dataset is comprised of data from the 2002-2008 NBA seasons. This is an unbalanced panel because the NBA had 29 teams for the first two years, until the Charlotte Bobcats were added as an expansion team in the 2004 season. A total of \( n=208 \) team observations were measured and recorded in the dataset.
Table 4.1. Summary Statistics Model One \( (n = 206 \text{ observations}) \)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definitions</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>WINPCT</td>
<td>Win Percentage in the regular season</td>
<td>0.5</td>
<td>0.15</td>
<td>0.16</td>
<td>0.82</td>
</tr>
<tr>
<td>SALARYRATIO</td>
<td>Team payroll divided by league average payroll</td>
<td>1</td>
<td>0.22</td>
<td>0.4</td>
<td>2.01</td>
</tr>
<tr>
<td>SALARYRATIOSQ</td>
<td>Square of Salaryratio</td>
<td>1.05</td>
<td>0.51</td>
<td>0.16</td>
<td>4.03</td>
</tr>
<tr>
<td>CAREERWINPCT</td>
<td>Coaches career win percentage</td>
<td>0.51</td>
<td>0.1</td>
<td>0.26</td>
<td>0.7</td>
</tr>
<tr>
<td>HHI</td>
<td>Herfindahl-Hirschman index for geographic diversity</td>
<td>0.77</td>
<td>0.17</td>
<td>0.34</td>
<td>1</td>
</tr>
<tr>
<td>International Minutes</td>
<td>International player minutes on team i</td>
<td>2297.22</td>
<td>2053.08</td>
<td>0</td>
<td>9263</td>
</tr>
</tbody>
</table>

Note: All data from Basketballreference.com, with the exception of salary information from Nba.com and eskimo.com/~pbender for years 2005-2006 and 206-2007.

The difference in player performance metrics of NBA players from different geographic regions is apparent when examining the test of means results displayed in Table 4.2. A difference in means test of player statistics by geographic region (using p-values) was performed on the original dataset. The results reveal that there are statistically significant differences in skill sets between players from different geographic regions.

Field goal percentage for all of the regions, with the exception of the ‘Other’ region, was seen as significantly different from the aggregated means of all other regions. The African region was seen as significantly different from the aggregate mean for all five of the performance measures. The European region (which had the second highest number of observations after the domestic region) had a statistically significant difference in mean from the domestic region in FGP and TOG. This could be the result of European training and coaching, which is known for its emphasis on shooting and skill work from a young age. Lazear highlights the necessity for disjointed skill sets in worker ability for hiring international talent, and Table 4.2 supports that notion (1999). However, Table 4.2 does not control for position or playing time so these differences may just as easily be due to those factors. In any event these differences in means deserve further investigation.
Table 4.2. Player Productivity Mean By Geographic Region

<table>
<thead>
<tr>
<th>Int’l Region</th>
<th>Performance Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FGP</td>
</tr>
<tr>
<td>Domestic (n=1495)</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.45</td>
</tr>
<tr>
<td>(P-value) to reject null</td>
<td>(0.000)</td>
</tr>
<tr>
<td>European (n=164)</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td><strong>0.46</strong></td>
</tr>
<tr>
<td>(P-value) to reject null</td>
<td>(0.093)</td>
</tr>
<tr>
<td>South American (n=52)</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.48</td>
</tr>
<tr>
<td>(P-value) to reject null</td>
<td>(0.006)</td>
</tr>
<tr>
<td>African (n=17)</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.48</td>
</tr>
<tr>
<td>(P-value) to reject null</td>
<td>(0.036)</td>
</tr>
<tr>
<td>Asia/AUS (n=12)</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.52</td>
</tr>
<tr>
<td>(P-value) to reject null</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Other (n=7)</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.48</td>
</tr>
<tr>
<td>(P-value) to reject null</td>
<td>(0.37)</td>
</tr>
</tbody>
</table>

Note: Mean values are pooled across players for all seasons measured with their weighted averages. Null hypothesis is that group identified mean is the same as all other observations in data set. P-values are assuming unequal variance. Bolded entries show significance at 5% to reject null hypothesis. Bolded and stars show 10% significance.

The low number of observations for certain geographic regions in Table 4.2 is also of concern. These low numbers lead to certain individual players contributing more to the mean performance measures than a domestic counterpart. For example, Yao Ming, a 7’5” Chinese center, is responsible for the inflated Asia/AUS RPG numbers as compared to the pooled means of other regions. This is because he accounts for the majority of observations in the years recorded, and always averaged many rebounds a game. This highlights the need to gather more years of data to increase the number of observations for both team and player statistics. Having established differences in worker talent by region we turn next to the estimation of the production function.

We employ ordinary least squares (OLS) with annual fixed effects (FE) and Newey-West robust standard errors because of the presence of second order serial correlation (detected by a LM test at the 5% significance level). A Jarque-Bera test confirms that the residuals are distributed normally at the 5% level. We do not employ team specific dummies because these would capture all team specific effects including the preponderance of a number of international players on a team. Two different models are estimated and the results are presented in Table 4.3.
Table 4.3. Regular Season Win Percentage

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Salary Ratio</td>
<td>0.37</td>
<td>0.20073</td>
</tr>
<tr>
<td></td>
<td>(0.13)</td>
<td>(0.09039)</td>
</tr>
<tr>
<td>Salary RatioSq</td>
<td>-0.14</td>
<td>-0.06794</td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td>(0.04381)</td>
</tr>
<tr>
<td>CareerWinPct</td>
<td>0.75</td>
<td>0.72191</td>
</tr>
<tr>
<td></td>
<td>(0.08)</td>
<td>(0.09473)</td>
</tr>
<tr>
<td>HHI</td>
<td>-0.12</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td></td>
</tr>
<tr>
<td>IntlMins</td>
<td>-</td>
<td>0.00001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.00000)</td>
</tr>
<tr>
<td>Overall R²</td>
<td>0.35</td>
<td>0.36</td>
</tr>
<tr>
<td>Observations</td>
<td>206</td>
<td>206</td>
</tr>
</tbody>
</table>

Note: Newey-West Robust standard errors in parentheses. Bold coefficients indicate significance at the 5% level or better.

In Model 1 after controlling for relative talent (via Salary Ratio and Salary RatioSq), coaching success (via CareerWinPct) we find that diversity (as measured by HHI) is negative and significant at the 5% significance level. This indicates that, ceteris paribus, teams with more diverse talent pools (smaller HHIs) have higher winning percentages. Model 2 replaces the HHI with the number of minutes played by international players. There is a positive and significant relationship between the number of international minutes played and the winning percentage of the team. The talent variable measured by (Salary Ratio and Salary Ratio Sq) and the coach’s career win percentage (CareerWinPct) are also significant in both models. The negative and significant coefficient on Salary Ratio Sq in both models supports a quadratic (concave down) relationship between the talent level and the win percentage of a team. In other words, increases in the team’s salary to the average league salary will buy more wins up to a point before diminishing returns set in.

5. **Conclusion**

Using annual NBA data from 2002-2008, we find that the regular season win percentage of a team is significantly impacted by the ratio of the team’s salary ratio to the league ratio, the coaches win percentage, the diversity of players and the number of minutes played by international players. This is because qualified
international players add to the scarce pool of homegrown talent thereby allowing more NBA franchises to field competitive teams. Fenn et. al find support for a similar hypothesis in the National Hockey League (Fenn et. al, 2005). Some NBA teams find it worthwhile to accommodate a diverse workforce because they can transcend regional differences and win games. In professional sports winning is the most important determinant of economic success. Winning in sports has been a powerful force for integration in various U.S. professional sports leagues. It appears that the influx of foreign players into the NBA is the latest example of the impact of international players on wins in a U.S. professional sports league.

REFERENCES


**Acknowledgements:** The authors wish to acknowledge outstanding research assistance from Annika Hansen.
ACCOUNTING HISTORY IN PERSPECTIVE: UNIFORM CPA EXAM TURNS 100

Darwin L. King
Carl J. Case
Kimberly M. Senecker
St. Bonaventure University

ABSTRACT
Professional accountancy examinations have been associated with the accounting profession for many years. They arrived with the advent of professional accounting associations. This paper provides a brief summary of the development of professional accounting examinations from their origin in Scotland and the first professional accounting association. In order to regulate entry into the accounting profession, the use of rigorous oral and written examinations were introduced. Professional accounting organizations quickly spread to England and Wales with associated examinations that evaluated the accounting knowledge and abilities of the applicant. Next, accounting associations or societies with their required examinations spread to Europe and the United States. This paper traces the evolution of professional examinations from the mid-1800’s to present day. Tremendous changes in the required professional accounting exams have taken place during that period. The story begins with accountancy exams required by the Institute of Chartered Accountants of Scotland (ICAS), which was the first recognized professional accounting association originating in 1854. This paper concludes with changes planned in 2017 for the U.S. Certified Public Accountants (CPA) Examination on its 100th anniversary.

Key words: Uniform CPA Exam, Chartered Accountants, first Scottish Society, Uniform CPA Exam in U.s in 1917

INTRODUCTION
The practice of accounting can be traced back nearly 6,000 years. Rudimentary accounting records were first discovered in approximately 4000 B.C. in the form of records of income from temples in lower Mesopotamia (ICAEW, 2015). More advanced accounting records, using the double entry format, were discovered in the “Massari Ledgers” of the Commune of Genoa of 1340. This was significantly earlier than 1494 when Luca Pacioli’s “Summa” was first published (Peragallo, 1956). Since little verifiable information remains concerning the Massari Ledgers, Franciscan Friar and mathematician Luca Pacioli continues to be recognized as the “Father of Accounting and Bookkeeping.” His 27 page
“Summa” was sold to merchants of the period as a reference text. It was also the first known text to contain algebra and the symbols for plus and minus. These symbols became standard notation for mathematicians during the Italian Renaissance period (Sangster, Stoner, McCarthy, 2007).

The practice of accountancy spread rapidly following Pacioli’s contributions with accounting texts becoming available in England. According to the Institute of Chartered Accountants in England and Wales (ICAEW), one of the earliest accounting societies known, an early text by Hugh Oldcastle printed in London in 1543 discussed the “Debitor and Creditor” relationships and explained methods to keep books in good order (ICAEW, 2015). Also, according to the ICAEW, Robert Colinson authored the first Scottish accounting text in 1683. He discussed, in his text, the art of bookkeeping using the Italian Method, which utilized the concept of debits and credits in accounting entries.

The practice of accountancy spread throughout the world prior to the creation of professional accounting associations, which typically required the passing of an examination or series of exams in order to gain membership to the group. In the late 1700’s in Scotland, business directories published in large cities listed individuals who practiced accounting. This occurred prior to the creation of any formal association or society (Brown, 1905). These accountants were, no doubt, the predecessors to Scottish accountants who created the first profession accounting society approximately fifty years later.

Since professional accountancy examinations evolved with the formation of professional accounting societies, a brief review of the first accounting societies is in order. The first professional accounting association was created in Scotland in 1854. These early accountants, in an effort to limit competition, created professional examinations. Scottish accountants, according to Lee, identified a major threat to their financial well-being (Lee, 1996).

Legislation proposed in London to base all Scottish insolvency practice on English legal provisions basically meant that only lawyers, not accountants, could administer Scottish insolvency cases. This encouraged Scottish accountants to unite, and between 1854 and 1856 they obtained public support and backing of Scottish members of Parliament to defeat the proposed legislation. During the period of 1854 to 1867, a number of professional accounting societies were formed. In turn, these organizations required the passing of professional examinations in order to being accepted into membership. These early examinations were typically extremely rigorous in order to limit the number of practicing accountants. By decreasing competition, practicing accountants of the period charged topmost rates for their services.
Professional accounting associations originated in Scotland in the mid-1800’s. With the advent of these organizations came the development of professional accounting examinations which would provide a method to verify the person’s knowledge of accounting and related business topics. The rigor of the exams determined the number of individuals accepted into the elite group of professional accountants.

Scottish accounting societies or associations originated in the 1850’s. The Institute of Accountants in Edinburgh and The Glasgow Institute of Accountants and Actuaries were formed in 1854. In July 1854, the Glasgow group petitioned Queen Victoria for a Royal Charter. The petition was signed by 49 Glasgow accountants arguing that theirs was a distinct respectable profession that was seeking a seal of professionalism (Alexander, 2002). The Edinburgh accountants filed its petition to Queen Victoria on October 23, 1854 (Royal Charter of 1854). The petition listed the duties of an accountant which included knowledge of the general principles of law including insolvency, bankruptcy, and property rights. An additional required skill of accountants included in the petition was the ability to provide evidence on financial matters to the courts. This could have possibly been the origin of forensic accounting skills which are common today.

The petition for Royal Charger also stressed the need for “a body corporate and politic” with the power to make rules and by-laws for the qualification and admission of members.” The petition asked the Queen for the group to be given the name of “The Society of Accountants in Edinburgh.” On December 11, 1854 the petition was registered and sealed. The Royal Charter gave the organization the right to set minimum qualifications for admission to the society. Hence, professional accounting examinations and professional accounting associations developed and evolved in close proximity. Accounting examinations during this period were typically conducted orally and records of questions asked are extremely difficult to locate. The Edinburgh group has undergone a number of name changes over the years and is currently known as the Institute of Chartered Accountants of Scotland (ICAS, 2016).

Due to the close proximity of Scotland and England, it didn’t take long for accounting associations to spread to England. Accountants in England during the mid-1800’s performed basic bookkeeping functions involving the purchase and sale of various assets. With the growth of limited liability companies and large scale manufacturing, the need for more technically proficient accountants increased. These more highly skilled accountants had to deal with the depreciation of assets, inventory valuation, and the recently issued “Companies” legislation (ICAEW, 2015a).
Five major English professional accounting organizations were created between 1870 and 1877. These groups, in an effort to show strength in number and solidarity, formed The Institute of Chartered Accountants in England and Wales (ICAEW) which was established by Royal Charter on May 11, 1880 (Perks, 1993). The five member associations included the Incorporated Society of Liverpool Accountants (1870), the Institute of Accountants in London (1870) renamed the Institute of Accountants in 1872, the Manchester Institute of Accountants (1871), the Society of Accountants in England (1872), and the Sheffield Institute of Accountants (1877). According to the ICAEW, the Institute of Accountants in London appointed a committee of examiners in 1871 to either approve or disapprove new applicants for admission to the profession. Candidates, during this period, were given oral interview examinations (viva voce) (ICAEW, 2015).

The ICAEW began with a total of 527 members that rapidly increased to 1,025 members in February 1881 (Woolf, 1912). In an effort to support high standards, the ICAEW required new members, in July 1882, to show knowledge and aptitude by passing a set of oral examinations (Chatfield, 1974). The three exams were named the preliminary, intermediate, and final examinations. In these early days, the ICAEW tested both the general and accounting (business) knowledge of all candidates seeking admission to the organization.

The preliminary examination’s content included subjects such as English composition, arithmetic, algebra, geography, history of England, elementary Latin, and two additional subjects from a group including Greek, French, German, physics, chemistry, animal physiology, electricity, geology, and higher mathematics. This exam had to be passed prior to moving to the next level of testing. The preliminary exam included only the above mentioned wide range of general knowledge topics. Accounting subjects were tested in the second examination.

The second or intermediate exam included bookkeeping, accounting, auditing, adjustment of partnership and executorship accounts, rights of trustees, liquidations, and receiverships. This exam included accounting and legal topics and had to be passed prior to moving to the third exam.

The final exam’s content included principles of the laws of bankruptcy, joint-stock companies, mercantile practices, and arbitration (Littleton, 1933). This exam was probably the most difficult for the early accountants due to the fact that it included difficult legal topics that were commonly handled by attorneys.

The ICAEW soon converted the exams from an oral to a written format. According to the ICAEW library, the following are examples of questions found on the preliminary or first exam from 1882 to 1886. Notice the wide range of subject areas included in the exam.
English Composition – Which of Shakespeare’s plays is, in your opinion, the most beautiful? Give your reasons, and the argument of the play.

Arithmetic – The diameter of a carriage wheel is 50 inches, and the circumference measures 3.14159 times the diameter, find the approximate number of times the wheel revolves in a second, if the carriage travels at the rate of 12 miles per hour.

Geometry – If a quadrilateral is inscribed in a circle, the opposite angles are equal to two right angles. Prove this.

History of England and Geography – (a) Describe a canoe voyage from the source to the mouth of the Thames, giving the names of the towns passed and the tributaries which join the river.
(b) Explain carefully the respective claims of Stephen and Matilda to the crown of England.
(c) Write out twenty towns, ten rivers, six states, and five mountains of India.

French – Translate into French (the question included a very long quote).

Chemistry – Sulphur is said to be Allotropic. Describe experiments showing how different forms of Sulphur may be obtained.

Animal Physiology – What is the evidence that different functions are performed by the anterior and posterior roots of the Spinal Nerves respectively? What are those functions?

Magnetism, Electricity, Light & Heat – How may the mean coefficient of a metallic rod be determined?

These are just a sample of the questions asked on the preliminary examination. Notice that none of these pertain directly to accounting but do require a detailed knowledge of a wide range of subjects. It was obvious that these examinations were used to limit the number of admission to the ICAEW and the practice of accounting.

In the next portion of this paper, the authors review professional accountancy examinations administered in the United States. The accounting profession in the U.S. was certainly influenced by Chartered Accountants from England who created many of America’s largest public accounting firms. This occurred very late in the nineteenth century.

UNITED STATES – HISTORY OF UNIFORM CPA EXAMS
According to Littleton’s survey of accountants practicing in 1850, there were only 19 such individuals listed in the city directories of New York, Chicago, and Philadelphia (Littleton, 1942). This was due, at least in part, to the fact that as late as the 1870’s it was typical for stockholder groups to visit corporate offices in an attempt to verify financial information (Previts & Merino, 1998). This was possible because, in 1875, there were very few manufacturing firms with assets exceeding ten million dollars (Newman, 1967). With the rapid growth of the size
of corporations during this period, practicing accountants were asked to perform various duties including preparing cases for arbitration or suit, detecting improper journal entries, locating errors in the accounting records, and discovering fraudulent transactions (Previts & Merino, 1998). According to Keister, the accountant was becoming the “foe of deceit and the champion of honesty” (Keister, 1896).

An interesting fact, during the late 1800’s, was the predominance of men in the finance and accounting professions. Clews, an investment banker of the period, reported that “As speculators, women hitherto have been utter failures. They do not thrive in the atmosphere of Wall Street, for they do not seem to have the mental qualities required to take in the varied points of the situation upon which success in speculation depends” (Clews, 1887). This attitude changed somewhat shortly after the turn of the century when several women were admitted to practice as certified public accountants (Previts & Merino, 1998). It was evident that chauvinism reigned in the accounting and finance professions during the latter part of the nineteenth century.

Chartered British accountants were regular visitors to the United States in the late 1800’s as they reviewed and audited the accounting records of railroads, breweries, and other large corporations in which their British clients had invested. According to Chatfield, four of the original “Big Eight” accounting firms in the United States were created by immigrant English Chartered Accountants (Chatfield, 1974). One of the first “Big Eight” firms in America was created by members of the Price, Waterhouse and Company arriving approximately in 1873 and opening a permanent office in 1890. In addition, Scottish accountants James Marwick and Arthur Young were founding members of major U.S. accounting firms that bore their names. These Scottish and English accountants supported the establishment of accounting associations, which they felt were both desirable and beneficial for the profession. With the creation of accounting associations came the professional examinations required in order to gain membership.

In the United States, according to Previts & Merino, the U.S. Constitution considers professional education and licensing to be a privilege of the state (Previts & Merino, 1998). The first professional accountancy organizations included the state societies of New York, Pennsylvania, and Illinois. These groups provided leadership to the profession until 1905. A national accountancy organization did not exist until after this date.

The New York State Society of Certified Public Accountants (NYSSCPA) was incorporated in 1897. An interesting fact is that New York State’s CPA law, passed on April 17, 1896 and signed into law by N.Y Governor Levi P. Morton, was created with the intent to exclude British Chartered Accountants stating that certificates would only be given to U.S. citizens or individuals intending to become a citizen soon.
During the years of 1896 and 1897, a total of 112 CPA certificates were awarded in NYS (Previts & Merino, 1998). The majority of them, 108, were granted on waiver by the NYS Board of Regents. A CPA certificate was granted during that period to individuals who could prove they had been in reputable practice as public accountants since January 1, 1890. The four other individuals took the actual CPA exam which was first given on December 15-16, 1896. Of these four applicants, three passed the exam. The typical American born CPA candidate of this period was approximately 30 years old and had only a high school equivalent education. It would not be for several decades that a college education was required to attain the CPA title (Previts & Merino, 1998). In the U.S., there has always been a close relationship between accountants and universities that taught accounting and finance. For example, in October 1900, New York University agreed to a NYSSCPA’s proposal to add a School of Commerce, Accounts, and Finance (Previts & Merino, 1998).

The Pennsylvania Institute of Public Accountants was incorporated in 1897 with a CPA law soon following in 1899 (Ross, 1940). Similar to the NYSSCPAs, a close relationship existed between the accounting organization and local educational institutions. The Pennsylvania Institute offered potential members night courses on accounting and auditing confirming its support for accounting education. In 1904, the University of Pennsylvania’s Wharton School of Finance agreed to assume leadership for the Pennsylvania Institute’s educational course offerings (Bennett, 1922).

The third major state accountancy organization was the Illinois Society of Public Accountants. Many Chartered British Accountants who could not earn a NYS CPA certificate, due to the requirement to be a U.S. citizen, moved to Illinois to acquire the CPA title. With the significant British presence in Illinois, there was an attempt to resist the passage of a formal CPA law (Ross, 1940). The Illinois Institute allowed Chartered Accountants, whether citizens or not, to become members. British impact on the Illinois Institute in 1903 was obvious since 21 of the 53 members were from the firm of Price Waterhouse (Wilkinson, 1904). The Illinois Institute quickly convinced Northwestern University to create a school of business and commerce again reflecting the close bond between accounting societies and institutions of higher learning in many states.

During the late 1800’s and early 1900’s, each state that had enacted a CPA exam was responsible for the creation of the exam for that state. In recent history, applicants from all 50 states have taken a uniform standardized examination. The first practitioner group to attempt national recognition in the U.S. was the American Association of Public Accountants (AAPA) which was formed in 1886 and incorporated in New York State on August 20, 1887 (Edwards, 1960). The AAPA, when it originated, did not require the passage of a professional examination to become a member. During this period, many accounting
professionals received a CPA certificate through a waiver system where they need only prove that they had practiced accountancy for a prescribed number of years. In the early 1900’s, some states passed regulations requiring the passage of examinations in order to receive a CPA certificate. According to Previts & Merino, the Chicago Herald reported that in Illinois, between the years 1903 and 1908, a total of 111 people took the CPA exam with only six receiving a passing grade (Previts & Merino, 1998). A more extreme situation occurred in New York State which had a failure rate of over 90% in 1916 with only three people out of 156 successfully completing the exam. This created turmoil for the profession as people such as James G. Cannon, president of the Fourth National Bank in NYC accused the profession of attempting to create a monopoly. Accusations of arbitrary grading was supported by the statement of a NYS examiner who, when grading an exam, stated that he would give “one-fifth for the correctness of the answer and four-fifths for its arrangement” (Suffern, 1909). Clearly, this examiner had little concern for the degree of knowledge and understanding of the candidate but instead was interested in appearance and organization. The AAPA, mentioned above, existed from 1887 to 1917 when the organization changed its name to the Institute of Public Accountants (IPA). One of the requirements for membership in the IPA was five years of practical experience and passage of the institute’s professional examination. The IPA’s name was changed again in 1957 to its current name of the American Institute of Certified Public Accountants (AICPA, 2015). Beginning in the early 1900’s, the accounting profession realized the value of both a quality education and accounting work experience. During the early 1900’s, the educational requirement incorporated into state CPA legislation was a “high school degree or its equivalent” (Previts & Merino, 1998). Over time that educational prerequisite has increased from 120 credit hours to the current 150 credit hour requisite. The renaming of the American Association of Public Accountants to the Institute of Public Accountants in 1917 introduced policies such as the setting of respectable admission standards, enforcing its own code of conduct, and promoting a common professional identity for accountants throughout the country (Chatfield, 1974). The same year, the IPA passed a set of eight operational rules. These rules involved ethical behavior and the use of honest business practices by accountants. For example, the acceptance of “kickbacks” in return for client referrals was forbidden. Another regulation required that members of the IPA would not engage in activities which were either incompatible or inconsistent with the practice of accountancy. A third policy example was that members of the IPA would not express an opinion on a set of financial statements until a supervisor or partner had first examined them (Chatfield, 1974). Ethical behavior was a major concern for these early professional accountants as it is today.
Origin of the Uniform CPA Exam - 1917

The uniform CPA exam was developed as an admissions requirement for membership in the AAPA. With the name change from AAPA to the IPA, the organization offered a uniform examination for use by state boards of accountancy across the country. These examinations served to further the cause of professional cohesion and verified the accountant’s ability to practice and earn membership in the IPA. Prior to 1917, each state had its own licensing examination which, no doubt, differed significantly from state to state. The 1917 uniform examination was a 13 hour two day test with 6 ½ hours of exam time each day. This uniform exam was utilized by only a few states in 1917 but the number gradually increased over the years until it was used by all of the states beginning in the early 1950’s.

The uniform CPA exam of 1917 was administered on June 14\textsuperscript{th} and 15\textsuperscript{th} (AICPA, 2015). It contained four sections including Auditing, Commercial Law, Accounting Theory and Practice-Part I, and Accounting Theory and Practice-Part II. The Auditing section was administered on June 14\textsuperscript{th} from 9am to 12:30pm and contained a total of ten questions. Applicants were required to answer all of the questions in this section. The Commercial Law portion was given on June 14\textsuperscript{th} from 1:30pm to 4:30pm and contained 16 questions. The candidate could select any 10 but not more than 10 questions.

The Accounting Theory and Practice–Part I was administered on June 15\textsuperscript{th} from 9am to 12:30 pm and contained a total of five questions. The candidate was required to answer the first two questions and then could choose one final question from the three remaining items. The fourth and final section, Accounting Theory and Practice-Part II, was given on June 15\textsuperscript{th} from 1:30 to 4:30 and contained eight questions. The Candidate was required to answer any six of the questions but no more than that number.

According to the AICPA, the following are a sample of the questions from the original 1917 examination from all four parts.

\textit{Auditing} –

1. What do you understand to be meant by a balance sheet audit? What is its scope?
2. State what you consider to be the most important special problems arising in one of the following classes of audits and how you would deal with such problems. (Stock brokers, Moving picture producers, Breweries, Clubs and institutions, Retail stores, Land Companies, Executorship accounts)
3. An inventory is submitted to you certified by the manager of the business. Mention some of the principal steps you would take to confirm the correctness of the inventory figure appearing in the balance sheet.

\textit{Commercial Law} –
1. What is the provision of the statute of frauds with respect to sales of goods, wares, and merchandise?
2. Define or describe void, voidable, and unenforceable contracts.
3. To what extent are dividends received by (a) an individual and (b) a corporation subject to payment of tax by the recipients under the income tax law of 1916?

Accounting Theory and Practice—Part I—
1. In a lengthy paragraph, numerous partnership business transactions were described. Following this, the student had to: Prepare trading and profit and loss accounts of each of the partners, indicating the final adjustment to be made in closing up the partnership. Second, show how the above final adjustment would be modified if partner A proved to have no assets or liabilities outside the partnership.
2. A second problem listed three very different balance sheets and information on a fourth corporation organized for the purpose of consolidating the first three companies. The candidate was required to prepare a consolidated balance sheet using additional information presented in the problem.

Accounting Theory and Practice—Part II—
1. What are organizational expenses? How are they to be treated in the accounts? At what point do expenses cease to be organizational expenses and become operational expenses?
2. Explain in full the theoretical difficulties in regard to each of the three commonly used methods of distributing overhead burden in cost accounting. Show how the appropriateness of each system may be affected by the nature of the business in which it is employed. Give briefly your views on the proper treatment of “Idle Time.”

These questions provide some insight into the material covered on the first uniform CPA exam in 1917. Over time more and more states adopted the uniform CPA exam but it was a slow process. It was not until 1952 that all states used the uniform exam provided by the Institute of Public Accountants (IPA) (not renamed the AICPA until 1957).

UNIFORM EXAMINATION CHANGES THROUGH 1996

The first uniform CPA examination given on June 14–15, 1917, discussed above, contained four sections and was a total of 13 hours in length. The exam gradually expanded into a two and one-half day 19.5 hour exam administered in five parts. The extended exam included Wednesday afternoon and all day Thursday and Friday. The Auditing, Business Law, and Accounting Theory exam segments were each 3.5 hours in length (10.5 hours total). The Accounting Practice Parts I and II were both 4.5 hours in length (9 hours total). The exam was administered twice per year on the first consecutive Wednesday, Thursday, and Friday of May and November. Prior to 1994, the use of calculators was forbidden (AICPA, 2015).
A major revision to the exam took place in 1994. The exam was reduced to a total of 15.5 hours over a two day period. The four new sections were Business Law and Professional Responsibilities, Auditing, Accounting and Reporting, and Financial Accounting and Reporting. This revised exam introduced an extended Auditing test which increased from 3.5 hours to 4.5 hours. The Business Law exam was shortened from 3.5 hours to 3.0 hours. The previous Practice and Theory sections were reduced significantly from 12.5 hours to eight hours in the form of two new Accounting and Reporting sections. According to the AICPA, this decreased the total time allocated to Practice and Theory from 64% to 52% of the exam. The increased length of the Auditing exam increased the time allocated to such topics from 18% on the old exam to 29% on the new version.

Prior to 1996, examination questions were published and used by candidates when studying for future exams. In 1996, the CPA exam was closed meaning that content material remained confidential so the high quality questions could be reused on future exams. The popularity of personal computers (PC’s) would significantly change the administering of CPA exams in 2004.

**CPA EXAM BECOMES COMPUTERIZED IN 2004**

Until the end of 2003, the CPA exam was a pencil and paper exam that was administered only twice per year (May & November). Beginning in April 2004, the CPA exam was administered by computer in Prometric testing centers. The exam was reduced to 14 hours and comprised four sections (AICPA, 2015). The four parts included Auditing and Attestation (AUD), Business Environment and Concepts (BEC), Financial Accounting and Reporting (FAR), and Regulation (REG). The computerized exam includes multiple choice questions, task based simulations, and written communication tasks. Task based simulations use real life work situations to test the candidate’s in-depth knowledge and skills in a specific subject area. The written communications tasks were found only in the BEC section of the exam requiring the candidate to respond in memo or letter format to a scenario presented in the test question.

The AUD section (4 hrs.) included 90 multiple choice questions and seven task based simulations. The BEC portion (3 hrs.) contained 72 multiple choice questions and three written communication tasks. The FAR section (4 hrs.) was comprised of 90 multiple choice questions and seven task based simulation. Finally, the REG (3 hrs.) segment included 72 multiple choice questions and six task based simulations. Candidates were allowed to take the parts in any order that he or she chose but all four had to be completed within 18 months. The exam was graded on a scale from 0 to 99 with 75 the minimum passing score.

With the advent of computerized testing, the exams became available to candidates for the majority of the year rather than only in May and November. According to the AICPA’s Candidate Bulletin, the exam is offered the first two months of each calendar quarter (AICPA, 2013). No exams are offered during the
third month of the each quarter (March, June, September, and December). Candidates can take any or all of the sections of the exam during a two month testing window. For example, any or all parts of the exam can be completed during a January/February or April/May window. Candidates, however, are not allowed to take the same exam over again during any one two month testing window. According to the AICPA, during the period from 2004 to 2009, the exam was administered over one million times (AICPA, 2015).

The AICPA publishes the passing rate percentages for the exam on a quarterly basis. A brief summary of CPA passing rates for 2006, 2011, and 2016 are shown below listed by exam section. The figures shown are the cumulative averages for the year.

<table>
<thead>
<tr>
<th>Section</th>
<th>AUD</th>
<th>BEC</th>
<th>FAR</th>
<th>REG</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006</td>
<td>44.0%</td>
<td>43.8%</td>
<td>44.5%</td>
<td>42.3%</td>
</tr>
<tr>
<td>2011</td>
<td>45.6%</td>
<td>47.1%</td>
<td>45.6%</td>
<td>44.2%</td>
</tr>
<tr>
<td>2016</td>
<td>46.4%</td>
<td>56.6%</td>
<td>46.4%</td>
<td>49.4% (figures for first three quarters)</td>
</tr>
</tbody>
</table>

An interesting fact over the last ten years is that the passing rate on all sections of the exam has continued to increase quite substantially. For example, candidates passing the Business Environment and Concepts (BEC) section increased from 43.8% in 2006 to 56.6% in 2016 (first three quarters). Likewise, the Regulation (REG) section saw an increase in the passing rate from 42.3% in 2006 to 49.4% ten years later in 2016. Is this something similar to “grade inflation” found in many colleges and universities? It is something that the AICPA and the National Association of State Boards of Accountancy (NASBA) are, no doubt, watching.

**REVISIONS TO THE CPA EXAM FOR 2016 AND 2017**

Minor cosmetic revisions to the exam were introduced during 2016. The AICPA announced that beginning January 4, 2016, the CPA exam would have a revised look in order to make it more intuitive and aid candidates in navigating through the exam. The 2016 exam contains subtle changes to the welcome screen, section introduction, and direction and break screens. During 2016 exams, students taking a break from the exam, consistent with prior years, lose that break time from the testing minutes allowed. Another cosmetic improvement was an updated exam timer display that is more visible and easier to read. Finally, in July 2016, the Document Review Simulation (DRS) was introduced in the Audit, Financial, and Regulation sections of the exam. The DRS is an innovative task-based simulation designed to increase exam authenticity by presenting a realistic scenario simulating tasks performed by a CPA such as reviewing memos, bank statements, and other accounting documents.

Major revisions are scheduled for the exam in April 2017 according to a May 2016 Journal of Accountancy article (Tysiac, 2016). The AICPA collects a practice analysis periodically in order to measure the knowledge and skills required
in a CPA in order to protect the public interest and meet the needs of state boards of accountancy. A practice analysis conducted in early 2014 found that advances in technology and the outsourcing of routine tasks increasingly required CPAs to use higher-order cognitive skills and professional skepticism when planning and reviewing the work of others. In effect, the revised 2017 exam will place less emphasis on remembering and understanding skills and more emphasis on higher level analysis and evaluation. This will be accomplished through the use of a larger number of task based simulations in the AUD, FAR, and REG sections. In addition, simulations will be added to the BEC section of the April 2017 exam for the first time. In order to accomplish this, the overall percentage of multiple choice questions on the exam will be reduced from the current 60% to about 50% on and after April 2017.

The April 2017 exam will increase from 14 hours to 16 hours with the testing time for the BEC and REG sections each increasing by one hour. Scores on the AUD, FAR, and REG sections will be determined through scoring multiple choice questions (50%) and task based simulations (50%). The final score of the BEC portion is determined 50% from multiple choice questions, 35% from task based simulations, and 15% from written communications. The AICPA expects that exam scores will continue to be provided to candidates with 20 days except for the first few testing windows when graders take additional time to analyze candidates and set passing standards for each part of the exam utilizing the new format.

A final change in the April 2017 exam is the introduction of a break period for each exam section. A 15 minute break will be offered in the middle of each exam with the minutes taken for a break not reducing allowed exam time. In the past, candidates taking a bathroom break lost that time from the exam period. Finally, CPA exam experts estimate that the exam will be revised and updated every two years verses every seven years as has been the case in the past. This will be done to remain current given the rapidly changing business environment.

CONCLUSIONS AND SUMMARY
As the Uniform CPA exam turns 100 in 2017, this paper provided a brief history of the development of professional accountancy exams. Originating in Scotland and England and closely associated with the development of professional accounting societies, professional accountancy examinations are now fairly standard in all countries recognizing the accounting profession. Candidates desiring to enter the profession must show his or her knowledge and skills in numerous accounting subject areas. The U.S. CPA exam continues to evolve with the changing needs of the profession. In future years, no doubt, the changing technological and business environments will demand a higher level skill set from all entry level accountants. Future CPA exams must be designed to verify the knowledge, abilities, and skills of the candidates taking the exam, so that the public will continue to be protected and accountancy remain a respected profession.
REFERENCES

AICPA (2015). History of the AICPA,
http://www.aicpa.org/About+the+AICPA/Understanding+the+Organization/History+of+the+AICPA.htm
Institute of Chartered Accountants of Scotland (ICAS), 2016. Who We Are, https://www.iacs.com/about-iacs/who-we-are
AN INVESTIGATION OF INVESTOR REACTION TO CORPORATE SOCIAL RESPONSIBILITY MOTIVATION AND ETHICAL POSITION IN AN ENVIRONMENTAL CONTEXT

Karen Y. Green
The University of Toledo
Melloney C. Simerly
Western Kentucky University

ABSTRACT
This study uses an experimental approach to examine whether the motivation behind corporate social responsibility (CSR), required or voluntary or the type of CSR, preventative or corrective has repercussions for investment decisions made by non-professional socially responsible investors. Further, the impact of preventative and corrective CSR is examined in the context of investor ethical position. Our experiment utilizes a between- and a within-subjects design. Participants consisting of 141 graduate students were randomly assigned to two experimental conditions manipulating investment returns for two firms. This study provides evidence that the motivation for CSR, whether required or voluntary, does not influence the investment decisions of socially responsible investors. However, we find that socially responsible investors are willing to forego profit maximization in favor of supporting firms that are motivated to prevent, rather than correct, negative CSR consequences.

Keywords: corporate social responsibility; socially responsible investors; ethical position

INTRODUCTION

Many studies suggest shareholders may act in the best interest of society (McWilliams and Siegel, 2001; Sparkes and Cowton, 2004); however, the association between socially responsible investing and corporate social responsibility (CSR) is not compelling. There are substantial numbers of socially responsible investors participating in the U.S. stock market (SRI, 2012) and the demand for the inclusion of CSR information in annual reports is on the rise (Gray, 1992). This has led to an increase in CSR reporting (Gray, 1993; Azone, Manzini, and Noci, 1996). Prior literature argues that firms can benefit from including CSR in their strategic plans (van de Ven, 2008) and acting on CSR issues as they emerge is essential. However, the extant literature does not address the importance of the motivation for dedicating resources to CSR or the type of CSR practiced by firms. This study addresses how these issues may affect investor decisions as well as the role of ethical position in these matters.

Specifically, we investigate how non-professional socially responsible investors react to varied CSR motivation: implicit (legally required) or explicit
Green and Simerly

(voluntary) (Matten and Moon, 2008). Previous research reveals that CSR information is incorporated into investor decisions (Holm and Rikhardsson 2008). When considering socially responsible investors in particular, their primary incentive may be altruism (Piliavin and Charng, 1990; Andreoni, 1990) rather than monetary gain. Therefore, we posit that incremental monetary gain may be irrelevant to socially responsible investors under certain circumstances. We test this assumption by comparing the likelihood of investment for socially responsible investors in firms that practice implicit and explicit CSR.

Prior literature suggests that socially responsible investors seek to support firms that make a positive contribution to society (Schueth, 2003). In addition to motivation, firms can engage in two different types of CSR: preventative or corrective. Preventative CSR occurs ex ante, or prior to negative CSR consequences. Corrective CSR takes place ex post, or is a reaction to the occurrence of a negative CSR event. Consequently, we examine whether the type of CSR (corrective or preventative) affects investment decisions of socially responsible investors. However, the purpose of socially responsible investment is to align investment decisions with ethical values (Schueth, 2003). Consequently, when investors make decisions involving both social responsibility information and investment return information, an ethical judgment is necessary. Accordingly, we examine the impact of corrective and preventative CSR in the context of the ethical position (EP) of socially responsible investors. EP consists of two dimensions that can be dichotomized to reveal four EP orientations: absolutist, exceptionist, subjectivist, and situationist (Forsyth, 1980).

Our results suggest that the motivation behind CSR (implicit or explicit) does not significantly influence the investment decision made by socially responsible investors. We also find support that all four of the EP orientations for socially responsible investors demonstrate a significant increase in the likelihood of investing in a firm practicing preventative CSR, when no monetary reward or penalty is anticipated. Additionally, when investors can expect a monetary reward for investing in a corrective CSR firm, as predicted, exceptionists do not change their likelihood of investing in the preventative CSR firm. However, contrary to our prediction, absolutists demonstrate an increase in the likelihood of investing in the preventative CSR firm. We also find that both subjectivists and situationists do not alter their likelihood of investing in the firm practicing preventative CSR, even when they can anticipate a higher return from the firm practicing corrective CSR.

This study adds important evidence to the CSR literature concerning how socially responsible investors react to firm decisions concerning stakeholder accountability. In addition, this research equips managers and regulators with information about the CSR information that socially responsible investors find relevant when making investment decisions. Regardless of their EP, socially responsible investors remain committed and in the case of the absolutist, increase their commitment to firms that practice preventative CSR. These investors demonstrate this behavior even when they can earn higher returns with a corrective CSR firm that responds to environmental concerns, although this response occurs after stakeholders are adversely affected. We highlight that some socially
responsible investors are driven by altruism and are willing to forfeit monetary gain in favor of supporting firms that are accountable for the prevention of negative consequences to the community in which they operate. Our evidence also provides valuable information to managers who set the CSR agenda by identifying how socially responsible investors react to preventative and corrective CSR action. For instance, managers may want to proactively respond to potentially negative CSR events in order to continue to attract socially responsible investors. In the remainder of this paper, we discuss theory and prior research supporting our hypotheses, followed by a description of the results as well as a discussion of the conclusions and implications of the study.

BACKGROUND AND HYPOTHESES DEVELOPMENT

Socially responsible investors contribute $3.07 trillion out of the $25.2 trillion invested in the United States. This represents one out of every eight investment dollars and emphasizes the significance of socially responsible investment capital (SRI, 2012). Furthermore, McWilliams and Siegel (2001) include stockholder demands as a part of the pressure a firm receives to allocate resources to CSR. In a comparison of socially responsible investors to conventional investors, McLachlan and Gardner (2004) document disparities regarding decision-making styles, moral perceptions and the influence of ethical issues. We seek to examine differences among socially responsible investors as a group, starting with the examination of the reaction of socially responsible investors to varied CSR motivation and types of CSR.

The motive behind a firm’s CSR agenda can be implicit or explicit (Matten and Moon, 2008). An explicit CSR action is one that is initiated voluntarily without legal requirement. The firm has volunteered to act in a socially responsible manner. Conversely, implicit CSR activities are motivated by legal regulation. These are not discretionary decisions. A company engages in implicit CSR in order to stay within the legal confines of the environment in which it operates (Matten and Moon, 2008). However, will the motivation behind CSR influence investment decisions for socially responsible investors?

Holm and Rikhardsson (2008) provide support that disclosure of CSR information influences investment decisions for both experienced and novice investors. Further, theories in social psychology (Piliavin and Charng, 1990) and economics (Andreoni, 1990) argue that altruistic and egocentric factors can lead individuals to forego the maximization of personal gain to benefit others. Mackey, Mackey, and Barney (2007) develop an analytic model, providing evidence that socially responsible investors seek to gain both monetary compensation and a feeling of satisfaction from knowing they are investing in firms that have a positive impact on society. Miller (1992) contends that socially responsible investors are those that want to align their investment decisions with their own political, religious, moral or ethical values. The goal of the socially responsible investor is to catalyze their funds and encourage companies to make a positive impact on the community in which they operate. We propose that the driving factors for CSR will be irrelevant to socially responsible investors because the reason for the
initiation of CSR action does not change the favorable outcome for stakeholders. Therefore, any positive CSR action, regardless of the motivation (implicit or explicit) will still meet the goal of the socially responsible investor by making a beneficial contribution to society. This leads to the following hypothesis:

H1: There will be no difference between the likelihood of investing in firms with implicit or explicit CSR motivation among socially responsible investors.

Carroll (1979: 500) defines social responsibility as “the economic, legal, ethical, and discretionary expectations that society has of organizations at a given point in time.” CSR consists of firm policies and practices that attempt to reverse social consequences of conducting business, and contributes to the society in which the firm operates (Matten and Moon, 2008). Prior research indicates that CSR activities fall into five main categories: environmental stewardship, social/community stewardship, economic development and profitability, stakeholder interaction, and voluntariness (Dahlsrud, 2008). Voluntariness refers to whether a firm undertakes voluntary or legally mandated CSR action. A recent study provides evidence that CSR is a concern for investors, however, the role of CSR information in investor decision making may be context specific. Wang et al. (2011) provide evidence that individual investors and institutional investors did not incorporate CSR information into their investment decisions prior to a food contamination incident. However, after the negative CSR event, the behavior of institutional investors was influenced by firm CSR performance.

Attention to CSR is not always anticipated or pre-planned by managers. Many companies engage in CSR to mitigate public dissention. For instance, in 2009 Toyota carried out their largest safety recall to date (Mitchell, 2009). There was speculation that several deaths occurred because Toyota vehicles accelerated unexpectedly. Yet, the 2009 recall was not carried out until after several denials of responsibility by Toyota. After numerous complaints were filed in the United States with the National Highway Traffic Safety Administration (NHTSA), the NHTSA issued a warning about the safety concerns for 3.8 million Toyota vehicles (Bensinger and Vartabedian, 2010). It was with this warning that Toyota issued the recall. On November 25, 2009, Toyota announced a detailed strategy to remedy the cause of the acceleration issue (Toyota, 2009). Toyota’s business practices affected stakeholders directly and the firm reacted ex post, in order to satisfy the public. This is an example of corrective CSR.

On the other hand, some companies use substantial resources to prevent harm to stakeholders. In 2012, Ford Motor Company issued a recall for 11,500 models of the 2013 Ford Escape due to a faulty fuel line that could have cracked and leaking fuel. Ford’s inspectors found that there was a faulty part before any major consumer incidents occurred (Aylesworth, 2012). Likewise, on February 17, 2012, Johnson and Johnson issued a voluntary recall of their Children’s Grape Liquid Tylenol products. This recall was implemented after discovering that consumers were not using the dosing syringe correctly (Krauskopf, 2012). The company did not agree to the recall because of adverse
health reactions. In fact, the possibility of a health problem occurring was remote (Singer, 2012). The response of Ford Motor Company and Johnson and Johnson are preventative CSR actions occurring *ex ante*, prior to stakeholders experiencing the negative consequences of the firm’s business practices.

During the 1970’s, social responsibility was largely a social responsiveness movement (Carroll, 1991). However, the focus then transformed to include proactive implementation of CSR action. Waddock and Smith (2000) highlight the importance of CSR to a company’s mission and vision. This mission will inevitably lead to the development of a business strategy. Wilson (1975) proposes four possible business strategies: reaction, defense, accommodation, and proaction. Therefore, CSR strategy can be thought of as reactive, defensive, accommodating, proactive, or some combination thereof. Preventative CSR would be analogous to a proactive CSR business strategy.

CSR and firm performance are complementary (Waddock and Smit, 2000). Additionally, performance is often included in prior literature as a defining component of CSR (Dahlsrud, 2008). Carroll (1991) states that for CSR to be legitimate, it must address the entire scope of a firm’s obligations, which include economic obligations for environmental issues. When Ford Motor Company issued the proactive recall for the 2013 Ford Escape before any major consumer incidents were reported, they experienced a stock price decrease of only 0.32%, over a one-day trading period (Aylesworth, 2012; Yahoo! Finance, 2012). When Johnson and Johnson issued the voluntary recall of their Children’s Grape Liquid Tylenol products before any health issues occurred, they experienced an increase in stock price of 0.18% (Singer, 2012; Yahoo! Finance, 2012). Alternatively, after Toyota denied responsibility for the unexpected acceleration issue in 2009 and was pressured into issuing a recall after numerous complaints were filed by the NHTSA, they experienced a stock price decrease of 1.95% after announcing a detailed strategy to remedy the situation on November 25, 2009 (Bensinger and Vartabedian, 2010; Yahoo! Finance, 2012). Then on November 29, 2009, Toyota issued a safety advisory, and the stock price subsequently decreased by another 5.21% (Yahoo! Finance, 2012). Both positive and negative signals concerning CSR issues can have stock market consequences. Alternatively, firm performance may influence investor decisions in the context of CSR information. Consequently, CSR information in conjunction with stock return information may make a difference in investment decisions, particularly for socially responsible investors. Further, these decisions may be influenced by the ethical position of the investor.

Firms are pressured by shareholders to allocate resources to CSR (McWilliams and Siegel, 2001; Sparkes and Cowton, 2004). Additionally, the market reacts negatively to reports of social irresponsibility and investors withdraw support from companies that fail to act in the best interest of society (Davidson and Worrell, 1988). Sparkes and Cowton (2004) point out that shareholder pressure for companies to behave in a socially responsible manner has steadily increased since the 1970’s. Moreover, there is significant pressure for companies to respond appropriately to CSR concerns expressed in shareholder proxies. Given this evidence along with Carroll’s (1979) definition of CSR that
includes economic and ethical components, we argue that social responsibility issues are perceived by investors to be both an economical and ethical matter. Therefore, when investors make decisions involving both social responsibility and investment return information, an ethical judgment is necessary.

Forsyth (1980) highlights the importance of ethical position (EP) in evaluating decisions and determines that ethical judgments are driven by an individual’s moral philosophy. He classifies individuals based on their acceptance or rejection of objective moral rules, and whether they view a decision as ethical, based on achieving positive consequences for all or if the result of a decision can be a combination of positive and negative repercussions (Forsyth, 1980). Those who are relatively idealistic follow objective and moral absolute rules. Conversely, more relativistic individuals believe that a decision is deemed ethically correct, depending on the circumstance. Accordingly, EP is based on two separate dimensions: relativism and idealism (Forsyth, 1980; Shaub, Finn, and Munter, 1993). These are not opposing ethical beliefs but rather a continuum on which each individual has a high or low orientation for each dimension (Forsyth 1980). Based on the relativism and idealism constructs, Forsyth (1980) posits four EP sub-dimensions: absolutist, exceptionist, subjectivist, and situationist (see Figure 1).

Forsyth and Berger (1982) argue that these four EP orientations predict differences in moral judgment. They discover differences in how each EP sub-dimension judges cheating. Thus, we employ the EP sub-dimension schema for socially responsible investors to determine whether moral judgments play a role in investment strategies in light of CSR information.

**FIGURE 1**

**Taxonomy table of Ethical Ideology**

<table>
<thead>
<tr>
<th></th>
<th>High Relativism</th>
<th>Low Relativism</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>High Idealism</strong></td>
<td><strong>Situationist</strong></td>
<td><strong>Absolutist</strong></td>
</tr>
<tr>
<td></td>
<td>Those who analyze each moral judgment within the context of each situation</td>
<td>Those who believe that an absolute universal code of values will have the best outcome</td>
</tr>
<tr>
<td><strong>Low Idealism</strong></td>
<td><strong>Subjectivist</strong></td>
<td><strong>Exceptionist</strong></td>
</tr>
<tr>
<td></td>
<td>Those who use personal values / perspectives as opposed to the a universal code of values</td>
<td>Those who typically use an absolute universal code of values; However, they are open to exceptions</td>
</tr>
</tbody>
</table>

Adopted from Forsyth (1980)

Extant scholarship validates the use of EP for research concerning financial and business decisions (Greenfield et al., 2008; Douglas and Wier, 2000; Fernanadez-Feijoo, 2013; and Douglas and Wier, 2005). However, some of the
results concerning the differentiating effects of EP are mixed. Marques and Azevedo-Pereira (2008) find that ethical decisions do not differ among the four ethical ideologies proposed by Forsyth (1980). However, Hartikainen and Torstila (2004) provide evidence to the contrary. They survey over 500 financial professionals and demonstrate that the four types of EP orientations respond differently to job related ethical dilemmas.

Absolutists believe in natural laws and reject actions based on consequences (Forsyth and Berger, 1982). These individuals believe decisions should be consistent with absolute moral rules (Forsyth and Pope, 1984). According to theory, absolutists would view a firm practicing preventative CSR as following the moral rule of ‘do no harm,’ and a firm practicing corrective CSR as deviating from this rule, since the purpose of corrective CSR action is to remedy an already harmful situation. Exceptionists believe that universal rules are a primary source for moral behavior; however, they understand that decisions producing some undesirable consequences may be necessary to achieve the best outcome.

Situationists are prone to analyze each scenario in the context of other events. If the CSR action has the best possible outcome for all, then it is determined to be the most ethical choice (Forsyth and Pope, 1984). Lastly, subjectivists are prone to using personal feelings when judging whether an event is ethical (Forsyth and Pope, 1984). They believe that judgments should be made individually and decisions should not be based on objective information. These investors are more likely to view a firm that practices preventative CSR as positive; although, they are much less optimistic than situationists. Subjectivists believe that some negative consequences are unavoidable (Forsyth and Pope, 1984). These investors will acknowledge the possibility of not being able to achieve humanitarian social good even if they invest in a firm practicing preventative CSR.

Stakeholder theory suggests a relationship between EP and investment decisions for firms that practice preventive or corrective CSR. The goals of a firm practicing preventative CSR are to uphold the firm’s underlying ethical values, identify potential impairments, and remedy any potential problems prior to harming stakeholders. Stakeholder theory contends that organizations are responsible not only to shareholders, but also to “a person or group that affects or is affected by the achievement of the organization’s objectives” (Freeman, 1994: 46).

According to EP theory (Forsyth and Berger, 1982), socially responsible investors who are absolutists and exceptionists demonstrate low relativism and place considerable emphasis on universal ethical principles, rather than circumstances. However, absolutists are highly idealistic and adhere to the moral rule of ‘do no harm.’ In making investment decisions between preventative and corrective CSR firms, socially responsible investors who are absolutists will likely view corrective CSR actions as damaging because stakeholders have already been harmed. They may be more apt to support the preventative CSR firm, even if the corrective CSR firm demonstrates potential for providing higher investment
returns. Exceptionists are much less idealistic and understand that some negative consequences may be necessary. They may not view corrective CSR as negatively as the absolutist. Nonetheless, the exceptionist will also view the preventative CSR firm as a more favorable investment option, even if the corrective CSR firm promises a higher return, because the firm practicing corrective CSR does not provide a superior moral alternative, only a superior monetary alternative. This leads to the following hypothesis:

H2a: When returns are lower for a firm that practices preventative CSR than that of a firm that practices corrective CSR, absolutists and exceptionists will not alter their likelihood of investment in the firm practicing preventative CSR.

More relativistic socially responsible investors, according to Forsyth and Berger (1982), will have a tendency to evaluate investment decisions in light of investment return rather than by the type of CSR. Forsyth and Berger’s (1982) EP theory suggests that a more relativistic investor’s preference for investing in firms practicing preventative CSR will decrease when monetary rewards favor firms that practice corrective CSR. This insinuates that subjectivists and situationists will alter their ethical view when a monetary reward is at stake. This line of reasoning leads to the following hypothesis regarding interaction between the EP orientations and expected returns:

H2b: When returns are lower for a firm that practices preventative CSR than that of a firm that practices corrective CSR, subjectivists and situationists will demonstrate a decrease in the likelihood of investing in the preventative CSR firm.

**METHODODOLOGY**

**Sample Surveyed**

Participants consist of graduate business students from a large mid-Atlantic University. Given their work experience and knowledge, graduate students are valid surrogates for non-professional investors (Libby, Bloomfield, and Nelson, 2002; Elliott W. B., Hodge, Kennedy, and Pronk, 2007). Of the 141 participants, 9 were deemed not socially responsible, 29 failed manipulation check questions, 1 had a serious omission, resulting in the 102 (72.3%) useable responses. Therefore, the analysis in the remainder of this study uses the 102 socially responsible participants who chose to invest in the preventative CSR firm and passed the manipulation checks. The average age of the sample is 28.1 years with a gender composition of 59% male and 41% female. The average work experience is about 6 years and 58.82% of the participants indicate they have participated in investing activities in the past.
Measurement Instrument and Task

Our experimental survey provided disclosure statements for participants to review from two different firms in the automotive industry, AUTO Incorporated and CARS Corporation. The respondents were asked to assume that both firms are publicly traded, have comparable market shares, and similar product reputations. The survey contained two disclosures stating that both firms had an issue with a chemical leak and would need to cease operations to make repairs. The respondents were provided with analyst forecasts for projected stock prices for each firm and were asked to select the firm in which they would invest and then rate their likelihood of investing in the selected firm on a scale of 1 (not likely) to 9 (very likely).

The respondents reviewed one disclosure statement for each firm that revealed a situation in which employees of the firm may or may not have been affected by a potential carcinogen. For the first firm, AUTO, the disclosure states that the organization is halting production to make repairs to stop a chemical leak. The example using AUTO represents preventative CSR. Their decision to make repairs is not the result of a negative signal, such as litigation or negative press. AUTO’s CSR action is intended to prevent any possible health hazard, although, no health risks have been linked to the environmental issue. The second firm, CARS, is also experiencing a chemical leak and must stop production to make repairs. The situation for the CARS firm has put stakeholders’ health at serious risk and the leaking chemical has a potential link to cancer. This represents corrective CSR action since the company has received a citation and litigation is expected.

Socially responsible investors concentrate investment activities on firms that focus on ‘what is right’ as opposed to ‘what is required’; consequently, we predict that socially responsible investors will view the firm that practices preventative CSR as the superior firm. Therefore, to determine whether the participant is a socially responsible investor, subjects were asked to identify whether they would invest the firm practicing preventative CSR (AUTO) or corrective CSR (CARS). Those who selected the CARS firm were determined not to be socially responsible investors and their responses were removed from the study.

To test hypothesis 1, we provide the respondents with two conditions in a between-subjects design. In one condition, we describe the disclosure as required by regulation and in the second condition we describe the disclosure as voluntary. To test hypotheses 2a and 2b, analyst expected returns is manipulated in a second between-subjects manipulation. Participants were randomly assigned to two conditions, one in which the stock market returns were expected to be the same for both firms or another in which the firm practicing corrective CSR has a higher expected return than the preventative CSR firm. The return information included in the instrument was based on the average return for an automotive Company in 2011. These data were obtained through CRSP (using variable PRC) and the higher return projection for the hypothetical CARS firm is one standard deviation above
the average for the auto company in 2011. The purpose of the difference in return projections between CARS and AUTO is to examine the relationship between an incremental monetary reward and EP in light of preventative versus corrective CSR.

EP is a latent construct based on two dimensions, relativism and idealism, and is assessed through Forsyth’s (1980) Ethical Position Questionnaire (EPQ). The EPQ consists of 20 items, 10 items to measure each dimension. The participants are asked to rate their level of agreement with each item on a 9-point Likert scale. Each individual has a relativism and idealism score that is the average of the related items for each dimension. Reliability for the relativism and idealism items have a Cronbach’s α of 0.85 and 0.86, respectively. As suggested by Forsyth, relativism and idealism scores are separated at their respective sample median to create indicators for each EP sub-dimension variables: absolutist, exceptionist, subjectivist, and situationist (Forsyth, n.d.).

RESULTS

Regression analysis is used to analyze whether implicit or explicit CSR motivation is important to socially responsible investors. Group comparison t-tests as well as regression analysis is used to examine the effects of each of Forsyth’s EP orientations on investment recommendations. The sample sizes for EP orientation is as follows: situationist = 23; absolutist = 28; subjectivist = 27; and exceptionist = 24. Means, standard deviations, and correlations for all variables are presented in Table 1. As expected, the EP orientations are highly correlated with the relativism and idealism constructs. Note that the correlations among the EP orientations are appropriately low.

TABLE 1

<table>
<thead>
<tr>
<th>Mean</th>
<th>SD</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Idealism</td>
<td>0.500</td>
<td>0.050</td>
<td>1</td>
<td>0.040</td>
<td>-0.479***</td>
<td>-0.648***</td>
<td>-0.148***</td>
<td>-0.515***</td>
<td>0.032</td>
<td>0.002</td>
</tr>
<tr>
<td>2. Relativism</td>
<td>0.490</td>
<td>0.050</td>
<td>-0.518</td>
<td>1</td>
<td>-0.491***</td>
<td>-0.564***</td>
<td>-0.314***</td>
<td>-0.408***</td>
<td>0.101*</td>
<td>-0.136</td>
</tr>
<tr>
<td>3. Situationist</td>
<td>0.225</td>
<td>0.042</td>
<td>0.490***</td>
<td>0.529***</td>
<td>1</td>
<td>-0.383***</td>
<td>-0.290***</td>
<td>-0.324***</td>
<td>0.139</td>
<td>-0.128</td>
</tr>
<tr>
<td>4. Absolutist</td>
<td>0.275</td>
<td>0.044</td>
<td>0.504***</td>
<td>-0.659***</td>
<td>-0.332***</td>
<td>1</td>
<td>-0.341***</td>
<td>-0.369***</td>
<td>0.195**</td>
<td>0.164</td>
</tr>
<tr>
<td>5. Exceptionist</td>
<td>0.235</td>
<td>0.042</td>
<td>-0.477***</td>
<td>-0.397***</td>
<td>-0.299***</td>
<td>-0.341***</td>
<td>1</td>
<td>-0.335***</td>
<td>-0.071</td>
<td>0.093</td>
</tr>
<tr>
<td>6. Subjectivist</td>
<td>0.265</td>
<td>0.044</td>
<td>-0.529***</td>
<td>0.481***</td>
<td>-0.324***</td>
<td>-0.569***</td>
<td>-0.333***</td>
<td>1</td>
<td>0.135</td>
<td>-0.078</td>
</tr>
<tr>
<td>7. Trust</td>
<td>0.480</td>
<td>0.059</td>
<td>0.007</td>
<td>0.208**</td>
<td>0.126</td>
<td>-0.505**</td>
<td>-0.071</td>
<td>0.135</td>
<td>1</td>
<td>-0.078</td>
</tr>
<tr>
<td>8. Implicit</td>
<td>0.510</td>
<td>0.052</td>
<td>0.021</td>
<td>-0.158</td>
<td>-0.128</td>
<td>0.164</td>
<td>0.095</td>
<td>-0.078*</td>
<td>-0.078</td>
<td>1</td>
</tr>
<tr>
<td>9. Explicit</td>
<td>0.863</td>
<td>1.623</td>
<td>0.168*</td>
<td>-0.066</td>
<td>0.073</td>
<td>0.094</td>
<td>0.014</td>
<td>-0.178</td>
<td>0.024</td>
<td>0.017</td>
</tr>
</tbody>
</table>

Note: Pearson’s R above the diagonal and Spearman’s rho is below the diagonal. *, **, and *** indicate significance at the 10%, 5% and 1% levels, respectively.

The first hypothesis predicts that there will be no difference in the likelihood to invest in firms with implicit (explicit) CSR motivation among socially responsible investors. The results of t-test group comparisons for the investment likelihood in the preventative CSR firm between explicit and implicit CSR motivation suggests that socially responsible investors will react statistically different regarding whether the CSR motivation is explicit (voluntary) or implicit.
(required) (*p*-value = 0.03) (Table 2, Panel A). The mean investment likelihood among socially responsible investors for the explicit CSR motivation is 6.28, while the mean investment likelihood for the implicit CSR motivation is 5.54. Initially, this suggests that H1 may not hold. However, these results fail to be robust. The regression analysis provides support for H1. The indicator variable ‘Implicit’ (*p*-value = 0.12), which is equal to one if the CSR is required and zero if voluntary, is not statistically significant. Hence, the investment decisions for socially responsible investors do not appear to be significantly influenced by CSR motivation (Table 2, Panel B).

**TABLE 2**

**Analyses for Importance of Implicit and Explicit CSR Motivation for Non-Professional Socially Responsible Investors**

**Panel A: Comparison T-test**

<table>
<thead>
<tr>
<th>Mean Implicit</th>
<th>Mean Explicit</th>
<th><em>p</em>-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Implicit vs. Explicit</td>
<td>5.54</td>
<td>6.28</td>
</tr>
</tbody>
</table>

**Panel B: Regression Analysis**

*Dependent Variable: Investment Likelihood in Firm with Implicit vs. Explicit CSR Motivation*

<table>
<thead>
<tr>
<th>Unstandardized Beta Coefficient</th>
<th>Standard Error</th>
<th>Standard Deviation</th>
<th><em>p</em>-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Implicit/Explicit <em>-0.53</em></td>
<td>0.35</td>
<td>-1.53</td>
<td>0.12</td>
</tr>
</tbody>
</table>

Implicit/Explicit - binary variable equal to 1 if disclosure is implicit (voluntary) and 0 if it is explicit (required by law).

**TABLE 3**

**Comparison T-tests and Regression Analysis for Ethical Position Sub-Dimensions**

**Panel A: Comparison T-tests**

<table>
<thead>
<tr>
<th>Orientation</th>
<th>Former Means</th>
<th>Later Means</th>
<th><em>p</em>-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absolutist vs. Exceptionalist</td>
<td>5.12</td>
<td>5.64</td>
<td>0.54</td>
</tr>
<tr>
<td>Subjectivist vs. Exceptionalist</td>
<td>6.37</td>
<td>5.64</td>
<td>0.85</td>
</tr>
<tr>
<td>Situationist vs. Absolutist</td>
<td>6.42</td>
<td>5.12</td>
<td>0.02</td>
</tr>
<tr>
<td>Situationist vs. Subjectivist</td>
<td>6.42</td>
<td>6.37</td>
<td>0.18</td>
</tr>
</tbody>
</table>
Panel B: Comparison T-tests with Interactions

<table>
<thead>
<tr>
<th>Orientation</th>
<th>Means of Orientation</th>
<th>Means of Interaction (Orientation*Return)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Situationist</td>
<td>6.42</td>
<td>6.43</td>
<td>0.10</td>
</tr>
<tr>
<td>Subjectivist</td>
<td>6.37</td>
<td>5.56</td>
<td>0.06</td>
</tr>
<tr>
<td>Absolutist</td>
<td>5.12</td>
<td>6.67</td>
<td>0.00</td>
</tr>
<tr>
<td>Exceptionist</td>
<td>5.64</td>
<td>6.30</td>
<td>0.01</td>
</tr>
</tbody>
</table>

Panel C: Regression Analysis

Dependent Variable: Investment Likelihood in the Firm that Practices Preventative CSR

<table>
<thead>
<tr>
<th>Variable</th>
<th>Unstandardized Beta Coefficient</th>
<th>Standard Error</th>
<th>Standard Deviation</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Situationist</td>
<td>6.42</td>
<td>0.82</td>
<td>7.83</td>
<td>0.00</td>
</tr>
<tr>
<td>Absolutist</td>
<td>5.12</td>
<td>0.78</td>
<td>6.56</td>
<td>0.00</td>
</tr>
<tr>
<td>Subjectivist</td>
<td>6.37</td>
<td>0.72</td>
<td>8.80</td>
<td>0.00</td>
</tr>
<tr>
<td>Exceptionist</td>
<td>5.64</td>
<td>0.80</td>
<td>7.03</td>
<td>0.00</td>
</tr>
<tr>
<td>Return*Situationist</td>
<td>0.24</td>
<td>0.70</td>
<td>0.35</td>
<td>0.73</td>
</tr>
<tr>
<td>Return*Absolutist</td>
<td>1.74</td>
<td>0.64</td>
<td>2.63</td>
<td>0.01</td>
</tr>
<tr>
<td>Return*Subjectivist</td>
<td>-0.66</td>
<td>0.65</td>
<td>-1.01</td>
<td>0.31</td>
</tr>
<tr>
<td>Return*Exceptionist</td>
<td>0.66</td>
<td>0.69</td>
<td>0.95</td>
<td>0.35</td>
</tr>
<tr>
<td>Trust</td>
<td>0.03</td>
<td>0.10</td>
<td>0.26</td>
<td>0.80</td>
</tr>
</tbody>
</table>

R²: 0.16
Adjusted R²: 0.08

Variable definitions:
- Implicit: Indicator variable equal to 1 if disclosure is implicit (voluntary) and 0 if it is explicit (required by law)
- Situationist: Indicator variable equal to 1 if participant is a situationist and 0 otherwise
- Absolutist: Indicator variable equal to 1 if participant is an absolutist and 0 otherwise
- Subjectivist: Indicator variable equal to 1 if participant is a subjectivist and 0 otherwise
- Exceptionist: Indicator variable equal to 1 if participant is an exceptionist and 0 otherwise
- Return: Indicator variable equal to 1 if return is lower for the Preventive CSR firm and 0 if the returns between the preventative and corrective CSR firms are equal.
- Trust: respondent’s rating on the level of trust regarding management disclosure statements. Rated 1 through 9, where 9 indicates highest level of trust in disclosures.

Two sample mean comparison tests were performed to identify whether participants with various ethical positions have different investment recommendations. When examining the four EP orientations, the analysis provides evidence that only one comparison, the situationist compared to the absolutist, have statistically different investment recommendations (p-value = 0.02). The remainder sample comparisons fail to provide evidence of differences for the mean of the investment likelihood among the EP orientations (Table 3, Panel A).

To test hypotheses 2, each EP orientation is examined to determine whether these investors will adjust their view of ethical behavior in light of a monetary reward. We compare the mean of each orientation when the returns for
both firms are the same to the mean of each sub-dimension when the return for the firm practicing corrective CSR is projected to be more favorable than the return for the firm practicing preventative CSR. Two sample mean group comparison tests are performed to identify whether participants with different EP orientations have varied investment preferences. The results in Table 3, Panel B show that when the projected return is lower for the preventative CSR firm, socially responsible investors change their investment recommendation.

A regression analysis examining the effects of the EP orientations in each EP sub-dimension has a positive and significant increase in investment likelihood with \( p \)-values \( \leq 0.01 \) (see Table 3, Panel C). Absolutists, those who follow a universal code of ethics, are the least likely to invest (\( \beta = 5.12 \)) in the preventative CSR firm. It seems that absolutists view the event as unethical despite the firm’s proactive response to remedy the situation. Exceptionists are the second least likely to invest in the firm that practices preventative CSR (\( \beta = 5.64 \)). These investors are similar to absolutists but understand that deviations from absolute moral rules are sometimes necessary to arrive at the best outcome. Exceptionists view the firm’s preventive response as an action in the best interest for stakeholders. On the other hand, exceptionists appear less likely to invest due to the hazardous event.

Situationists have the highest investment recommendation (\( \beta = 6.42 \)), indicating that this type of investor uses the context of the scenario to determine whether the action is ethical. The other more relativistic EP, subjectivists, who determine whether a behavior is ethical based on own their personal values, have a similar investment recommendation (\( \beta = 6.37 \)). These socially conscience investors may view the preventative CSR action as the firm’s attempt to remedy the situation and increase safety for the stakeholders affected.

The results presented in Panel C of Table 3 demonstrate that H2a is partially supported. When the returns of preventative CSR firms are lower than the returns of corrective CSR firms, the likelihood of investment does not change for exceptionists, and absolutists show an increase in the likelihood of investing in preventative CSR firms, under the same conditions. This indicates that absolutists may reward preventative CSR firms even when faced with an alternative and more attractive investment in a corrective CSR firm. Alternatively, our results fail to support H2b. The interactions for returns and situationist, as well as returns and subjectivist, are not significant. This suggests that both subjectivists and situationists are unlikely to alter their view of what is deemed ethical when a pecuniary reward is provided. It may be the case that socially responsible investors are not easily influenced by monetary rewards and consequently, a larger monetary incentive is necessary to convince these investors to adjust their ethical threshold.

**CONCLUSION**

This study provides evidence that various aspects of CSR influence socially responsible non-professional investor recommendations. These investors are more likely to forgive a firm’s actions if the firm is proactive and upholds the ethical foundation of being a socially responsible. These investors also seem willing to forgo excess monetary returns in order to support firms that proactively
work to prevent stakeholder harm, regardless of whether the CSR motivation is explicit (voluntary) or implicit (required). Further, absolutists increase their investment recommendation in firms practicing preventative CSR, despite the expectation of a lower expected return. These results are congruent with Webley, Lewis, and Mackenzie (2001) who use an economic experiment to document that ethically focused investors are committed to ethical organizations and will continue to invest even if the organization performs poorly. Our study also lends empirical support for the argument provided by Winnett and Lewis (2000: 325) “… that the extra-economic (on the usual definition) motive in ethical investment somehow helps to justify their peculiar efficacy.”

Although capital markets have been predominately a means for larger firms to raise equity, the implementation of crowdfunding will grant smaller firms opportunities to raise capital from individual investors. With this possibility, stemming from the Jumpstart our Business Startups Act (JOBS Act), firms will be able to raise up to one million dollars in capital via online resources from unaccredited investors (U.S. Securities and Exchange Commission, 2012). Moreover, the JOBS Act requires fewer financial disclosures. Accordingly, it will be important for small startup businesses to be knowledgeable about investor behavior. If a startup firm practices preventative CSR for issues concerning the environment and provides disclosure information about such practices, socially responsible investors may accept a smaller return in favor of investing in a firm that takes action to avoid negative consequences to the society in which it operates. In other words, if these businesses are proactively addressing adverse environmental effects, socially responsible investors are more apt to continue to offer their support.

This study also substantiates that some socially responsible investors do not behave in a self-serving manner in pursuit of only profit maximization. Socially responsible investors support firms that take action in line with their own ethical values and beliefs. The results here support the notion that these investors are driven by altruism. They are willing to forego profit in favor of supporting a firm that adheres to a strong ethical foundation.

CSR is a commonly used by organizations to increase transparency (Fernandez-Feijoo, Romero, and Ruiz, 2013). Nevertheless, investors are mindful of organizations’ CSR disclosure statements (Besio and Pronzini, 2013). Employees are also responsive to their organization’s CSR practices, particularly whether these practices are voluntary or required (Alfaro-Barrantes, 2012). Future research could examine the relationship between employee relations and firm CSR motivation concerning the environment. Furthermore, this study examines the investing choice between two organizations with an expected return difference of one standard deviation between firms that practice preventative and corrective environmental CSR. By increasing the monetary reward, researchers could determine at what threshold do socially responsible investors adjust their
investment recommendations and choose the monetary reward rather than support a firm practicing preventative CSR for environmental concerns.

REFERENCES


SEASONALITY IN THE MONTHLY RETURNS OF LARGE STOCKS: 1926 TO 2013

Shaikh A. Hamid
Southern New Hampshire University

ABSTRACT

This paper looks at the seasonality in large stocks from the perspective of means and standard deviations of monthly returns from 1926-2013, seasonality in the case of four sub-periods carved based on structural changes in the economy, differences in the returns and standard deviations between Republican presidential versus Democratic presidential months, and seasonality during these two types of presidencies. We have explored and found monthly seasonality for the entire data period: 1926-2013, and also for the four sub-periods. We find a significantly negative September effect in the monthly returns for the entire data period which appears to be largely the result of the monthly returns of September in the fourth sub-period (1993 to 2013). For the entire data series, the volatility of the monthly returns of April was significantly higher compared to the volatility of returns of the other eleven months stacked. We find significantly lower volatility for three other months (January, February, and December). We also find greater volatility of monthly returns in October compared to the volatility of the returns of other eleven months in the last two sub-periods – significant only in the third sub-period. We look at the performance of large stocks during Republican and Democratic presidential months and find interesting results. Mean returns have been higher during Democratic months compared to Republican months at lower risk – contrary to conventional perception.

Key Words: Monthly Seasonality, Large Stocks, Volatility, Republican Months, Democratic Months

INTRODUCTION

We look at the monthly seasonality in large stock returns over the period January 1926 to December 2013. Existence of seasonality will provide guideline on strategy for investing in large stocks. We look at the seasonality in monthly returns of large stocks, firstly, for the whole data period, and then for sub-periods so as to get deeper insight. We also explore seasonality in monthly returns under Republican and Democratic presidential months. Further we explore if there are differences in the average monthly returns of large stocks under Republican versus Democratic presidencies for the entire data sample.
LITERATURE SURVEY

Since the time stock exchanges were first established, traders and investors have exhaustively looked for patterns in stock prices that they could exploit to realize superior returns. As early as 1900, Bachelier characterized security prices as being efficient. Over thirty years later came the landmark work by Cowles (1933) in which he documented the inability of forty-five professional agencies to forecast stock prices. The conclusion was that stock prices are random—in general they do not exhibit patterns. This was followed by the researches of Working (1934), Cowles and Jones (1937), Kendall (1953), and Osborne (1959, 1962). They documented that stock and commodity prices behave like a random walk—as if they are independent random drawings. These empirical works were buttressed by the theoretical work of Samuelson (1965) and Mandelbrot (1966). Fama (1965) also contributed to this body of literature which came to be termed the ‘random walk hypothesis.’ In 1970, Fama came up with the ‘efficient markets hypothesis’ (EMH). [In economics, Muth (1961) developed this hypothesis independently which was termed rational expectations hypothesis.] This hypothesis postulates that stock prices reflect all available information; they change in response to new information; since new information by definition cannot be deduced from previous information, new information must be independent over time; if the arrival of new information is random, stock price changes are random—the changes cannot be anticipated; hence it is not possible to generate risk-adjusted abnormal returns from stocks. Bernstein (1992) provides an overview of the developments of EMH.

The overall finding is that it is difficult to earn above-average profits by trading on publicly available information because it is already incorporated in securities prices. However, some researchers have been able to identify profitable opportunities or anomalies. These findings go against the concept of efficient markets. As a result, some academics have deprecated the concept. The adherents of the new camp may possibly be increasing. Among the various anomalies discovered, the January effect is possibly the most well-known. It has been documented for financial markets across the globe. The first evidence of returns in January exceeding those of other months comes from Wachtel (1942). After thirty-three years, Officer (1975) presented further evidence followed by Rozell and Kinney (1976). Wachtel introduced the concept of January effect in 1942, but Rozell and Kinney’s article in the widely respected Journal of Financial Economics was the first evidence of January effect that attracted widespread attention. These findings challenged the concept of efficient markets hypothesis that securities markets reflect all available information and hence it is not possible to garner positive risk-adjusted returns.

Reinganum (1983) has advanced the hypothesis that January experiences rebound in stock prices after tax-loss selling that is undertaken in December. The hypothesis is that before the end of the tax year, people sell stocks that have declined in price during the previous months so they may realize the capital losses; these investors put back the proceeds into the market in January; the higher demand
for stocks push stock prices up creating the January effect. Reinganum found that within firm size classes, firms for which price decline was more pronounced had larger January returns. Ritter (1988) has documented that the ratio of stock purchases to sales of individual investors hits an annual low at the end of December and an annual high at the beginning of January.

Haugen and Lakonishok (1988) have advanced the hypothesis that the January effect is a result of simultaneous reentry into aggressive investment strategy by professional fund managers who have parked money in their performance benchmarks so as to lock in their investment performance during the previous year.

A major finding that comes out of the researches is the size effect: small-capitalization firms earn higher returns than large-capitalization firms. Banz (1981) and Reinganum (1981) were the first researchers to discover the small-firm effect. Their finding was supported by Brown, Keim, Kelidon and Marsh (1983), Kato and Schallheim (1985), Fama and French (1992), Berk (1995), Baker and Limmack (1998), and Garza-Gromez, Hodoshima and Kunimura (1998). Keim (1983), Reinganum (1983), Blume and Stambaugh (1983) and Roll (1983) find that majority of the return of small-capitalization stocks occurs in January -- in the first two weeks of the month. This phenomenon came to be known as the small-firm-in-January effect. Keim found that small firms outperformed large firms in every year from 1963 to 1979.

It has been argued that the January effect is most pronounced for the smaller firms because the small firms are more volatile and more prone to price declines and hence more subject to tax-loss selling.

Arbel and Strebel (1983) found that the January effect was largest for firms neglected by institutional investors. This was termed the neglected-firm effect. The hypothesis is that small firms tend to be neglected by large institutional traders; this causes information deficiency which makes them riskier prompting investors to require higher returns.

Haugen and Jorion (1996) use data for the stocks in New York Stock Exchange from 1926 to 1993 and find that for smaller stocks January returns are significantly larger than for other months. This work also indicates that excess returns in January may be declining in later years. This was also found by Riepe (1998, 2001).

Agrawal and Tandon (1994) find for nineteen countries covering data for 1970’s and 1980’s that the mean January returns are high – significantly high for eleven countries. Hawawini and Keim (2000) survey international findings and show that the high returns for January relative to other months, if used as explanatory variable, better accounts for cross-sectional returns of stocks than the CAPM beta or some other data-driven models proposed in recent times. Giovanis (2009) look at calendar effects in fifty-five stock market indices.

A large number of researchers have explored monthly seasonality for various countries and regions that include Australia, Africa, China, India, Kuwait, Lithuania, Macedonia, Malaysia, Mauritius, Pacific-Basin countries (Japan, Korea, Taiwan, Hong Kong, Singapore), Spain, South-East Asia, UK, and
Vietnam. For brevity, the detailed references are not provided. Angelovska (1914) and Friday and Hoang (2015) include in their references many of these studies.

We intend to contribute to this growing literature by comparing the behavior of large stocks and how the behavior changed over time. The next section describes the descriptive statistics of large stocks, followed by results of analyses of monthly seasonality, pattern of returns during Republican and Democratic presidential months, behavior of returns during expansion and recession, and average returns during wars and financial crises. We also look at annualized returns. Some tentative conclusions follow.

DATA SOURCE AND METHODOLOGY
Monthly returns data for large stocks are taken from “Stocks, Bonds, Bills and Inflation Yearbook 2014”. The data is broken up into returns in the form of capital appreciation, and income returns. The returns are for S&P 500 stocks. Large stocks are represented by S&P 500 Composite with dividend reinvestment. It is S&P 500 stocks from 1957 to the present; from 1926 to 1956, it is S&P 90.

HYPOTHESES
We study the seasonality in terms of monthly total returns of large stocks from three approaches:

1. If the mean of monthly total returns is different from zero for the sample as well as for each month in the sample. We subject the mean return for a given month \( i \) to the following hypothesis test: \( H_0: \mu_i = 0 \) vs. \( H_a: \mu_i \neq 0 \).

2. If the mean of the monthly total returns for a month is significantly different from the mean of the other 11 months stacked. We conduct the following hypothesis test for a given month \( i \): \( H_0: \mu_i = \mu_j \) vs. \( H_a: \mu_i \neq \mu_j \), where \( j = \{1, 2, \ldots, i-1, i+1, \ldots, 11, 12\} \). Since the variances for the periods \( i \) and \( j \) were unequal in many cases, we used the more conservative t-test assuming unequal variances.

3. If the variance of the monthly total returns for a given month is different from the variance of the other eleven months stacked. We conduct the following hypothesis test for a given month \( i \): \( H_0: \sigma_i^2 = \sigma_j^2 \) vs. \( H_a: \sigma_i^2 \neq \sigma_j^2 \), where \( j = \{1, 2, \ldots, i-1, i+1, \ldots, 11, 12\} \).

Seasonality has been looked at in terms of mean of returns. Our research tells us no research has looked at seasonality in terms of variance of returns. If the markets are efficient, variance of the monthly returns should not be significantly different.

In addition to standard t-test which assumes normal distribution of the data, we also use Kruskal-Wallis non-parametric test which tests for differences among several population medians, and does not depend on normal distribution of data. We also use Mood’s Median Test which performs a nonparametric analysis of a one-way layout. Further, we use Mann-Whitney test which performs a two-sample Wilcoxon rank sum of the equality of two population medians. All these
non-parametric tests are robust in the presence of outliers, as is the case with financial markets data.

Many studies have used the dummy variable methodology to detect seasonality in market variables. Chien, Lee and Wang (2002) provide statistical analysis and empirical evidence that the methodology may provide misleading results. We avoid this methodology and use the methodology developed by Hamid and Dhakar (2008). Unless otherwise stated, significance is at 5% level.

In addition to analyzing the data for the entire period (January 1926 to December 2013), to gain deeper insight on seasonality of large stocks we divide the period into the following sub-periods:

- January 1926 to December 1945 (includes the Great Depression years and World War II);
- January 1946 to December 1972 (includes the stable period after World War II and the Breton Woods fixed exchange rate era, and ending with the breakdown of that era in 1972);
- January 1973 to December 1992 (includes the volatile world after the first oil crisis of 1973 and the post-Breton Woods era);
- January 1993 to December 2013 (includes the run-up to the dotcom boom, the subsequent bust, and the Great Recession after 2008).

We find that the month effect is sensitive to the time period under study.

**ANALYSIS OF RESULTS**

We test the three hypotheses above using first the full data set (1926 to 2013) and then the data of for the four sub-periods.

**Full Sample: 1926-2013**

The 88-year period consists of 1,056 months. The mean of monthly returns for the entire period is 0.95% (Table 1) and is significantly greater than zero. The mean monthly return of July is the highest (1.79%), followed by December (1.71%), April (1.53%), November (1.40%), and January (1.37%). All of these returns are significantly greater than zero. But none of these returns are significantly greater than the mean returns of the other eleven months stacked. This finding pretty closely follows the means of monthly changes found by Hamid and Dhakar (2016) for Dow Jones Industrial Average for 1896 to 2013 and by Lakonishok and Smidt (1988) for the period 1896 to 1986. The DJIA consists of stocks that are among the largest.
Table 1: Descriptive Statistics and Monthly Seasonality: 1926-2013

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>0.95</td>
<td>1.37</td>
<td>0.72</td>
<td>0.53</td>
<td>0.51</td>
<td>0.98</td>
<td>0.79</td>
<td>0.29</td>
<td>0.76</td>
<td>0.50</td>
</tr>
<tr>
<td>median</td>
<td>0.88</td>
<td>0.88</td>
<td>0.88</td>
<td>0.88</td>
<td>0.88</td>
<td>0.88</td>
<td>0.88</td>
<td>0.88</td>
<td>0.88</td>
<td>0.88</td>
</tr>
<tr>
<td>st Dev</td>
<td>0.48</td>
<td>0.73</td>
<td>0.22</td>
<td>0.05</td>
<td>0.64</td>
<td>0.66</td>
<td>0.31</td>
<td>0.13</td>
<td>0.18</td>
<td>0.97</td>
</tr>
</tbody>
</table>

Note: 1. “Pos” implies the mean of monthly returns was significantly greater than zero; “Neg” implies the mean of monthly returns was significantly less than zero; “Higher” implies the mean or the variance of monthly returns for a month was significantly higher than those of the other eleven months stacked; “Lower” implies the mean or the variance of monthly returns for a month was significantly lower than those of the other eleven months stacked. Numbers are rounded in all tables to two decimal places.

2. The results of tests shown in Table 3 onwards are rounded to two decimal places fit page width.

The mean return for September (-0.76%) is significantly lower than the mean return of the other eleven months stacked. However, December’s mean return is significantly different from the mean return of the other eleven months stacked for a p value of 0.052. So we have a negative September effect for large stocks which was also found by Hamid and Dhakar (2016) for the DJIA. We also find variances of monthly returns for April being significantly higher than that of the other eleven months stacked, and we find significantly lower variances for January, February, and December. So we find seasonality in terms of mean and also in terms of variance. Both these findings go against the notion of efficient markets.

Figure 1 shows the cyclicity of monthly means of large stock returns and the dip in the mean of the monthly returns from July to September, and the subsequent increase until December. Short term traders stand to gain by shorting in July and buying back in September, or buying in September and selling in December.
First Sub-period: 1926-1945

The mean monthly return for the first sub-period (Table 2) was 0.93% -- significantly different from zero at a p value of 0.09. This sub-period contained the Great Depression era when returns were rather depressed. The mean monthly return for August was the highest (4.59%) followed by July (4.02%), and June (3.64%). The mean return of March (-2.07%) was the lowest.

TABLE 2: Descriptive Statistics and Monthly Seasonality: 1926-1945

<table>
<thead>
<tr>
<th>Month</th>
<th>Ill</th>
<th>an</th>
<th>eb</th>
<th>ar</th>
<th>pr</th>
<th>ay</th>
<th>un</th>
<th>ul</th>
<th>ug</th>
<th>ep</th>
<th>ct</th>
<th>ov</th>
<th>ec</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>1.37</td>
<td>0.37</td>
<td>0.72</td>
<td>1.53</td>
<td>0.51</td>
<td>0.98</td>
<td>1.79</td>
<td>1.29</td>
<td>0.76</td>
<td>0.50</td>
<td>1.40</td>
<td>1.71</td>
<td></td>
</tr>
<tr>
<td>Stan. Dev</td>
<td>0.93</td>
<td>0.47</td>
<td>0.15</td>
<td>2.07</td>
<td>0.31</td>
<td>0.36</td>
<td>0.64</td>
<td>0.02</td>
<td>0.59</td>
<td>1.75</td>
<td>1.57</td>
<td>-0.10</td>
<td>-0.67</td>
</tr>
<tr>
<td>Median</td>
<td>0.58</td>
<td>0.55</td>
<td>0.62</td>
<td>0.06</td>
<td>0.26</td>
<td>0.89</td>
<td>0.81</td>
<td>0.29</td>
<td>0.15</td>
<td>0.77</td>
<td>1.16</td>
<td>0.55</td>
<td>0.25</td>
</tr>
<tr>
<td>t-value (p=0)</td>
<td>0.93</td>
<td>0.47</td>
<td>0.15</td>
<td>2.07</td>
<td>0.31</td>
<td>0.36</td>
<td>0.64</td>
<td>0.02</td>
<td>0.59</td>
<td>1.75</td>
<td>1.57</td>
<td>-0.10</td>
<td>-0.67</td>
</tr>
<tr>
<td>t-value (t-test)</td>
<td>0.93</td>
<td>0.47</td>
<td>0.15</td>
<td>2.07</td>
<td>0.31</td>
<td>0.36</td>
<td>0.64</td>
<td>0.02</td>
<td>0.59</td>
<td>1.75</td>
<td>1.57</td>
<td>-0.10</td>
<td>-0.67</td>
</tr>
<tr>
<td>F-value (F-test)</td>
<td>0.93</td>
<td>0.47</td>
<td>0.15</td>
<td>2.07</td>
<td>0.31</td>
<td>0.36</td>
<td>0.64</td>
<td>0.02</td>
<td>0.59</td>
<td>1.75</td>
<td>1.57</td>
<td>-0.10</td>
<td>-0.67</td>
</tr>
<tr>
<td>Mean Return Effect (Mean)</td>
<td>0.93</td>
<td>0.47</td>
<td>0.15</td>
<td>2.07</td>
<td>0.31</td>
<td>0.36</td>
<td>0.64</td>
<td>0.02</td>
<td>0.59</td>
<td>1.75</td>
<td>1.57</td>
<td>-0.10</td>
<td>-0.67</td>
</tr>
<tr>
<td>Mean Return Effect (Var)</td>
<td>0.93</td>
<td>0.47</td>
<td>0.15</td>
<td>2.07</td>
<td>0.31</td>
<td>0.36</td>
<td>0.64</td>
<td>0.02</td>
<td>0.59</td>
<td>1.75</td>
<td>1.57</td>
<td>-0.10</td>
<td>-0.67</td>
</tr>
</tbody>
</table>

Note: See note under Table 3; numbers are rounded to two decimal places.
However, presumably because there are only 20 data points per month, only the mean of August is significant at a p value of 0.04. The returns of June and July are significant at a p value of 0.08 and 0.09 respectively. The mean return of none of the months is significantly different at 5% level from those of the other eleven months stacked. August’s mean was different from the mean of the other eleven months stacked at a p value of 0.08. The mean return of March was negative and different from the mean of returns of the other eleven months stacked at a p value of 0.10. The variance of the returns of January, February and December were significantly lower compared to variance of returns of the other months stacked, whereas April saw higher variance compared to the other eleven months. April also had the widest range among all the months (62.53%). The standard deviations of this sub-period ranges from about 5% to 12.30% and are overall much higher compared to those of the other three sub-periods. So this period, characterized by the Great Depression and the Second World War exhibited greater volatility of monthly returns. The markets of recent times appear more volatile. But that is in terms of standard deviation of daily returns and not in terms of standard deviation of monthly returns [Hamid and Dhakar (2005)]. Kruskal-Wallis nonparametric test of difference in the medians of the twelve months shows no significant difference (H = 14.12, p = 0.23). Mood Median test also shows no significant difference in the medians of the twelve months (Chi-square = 5.78, p = 0.28).

**Second Sub-period: 1946-1972**

In the post-second World War period of fixed exchange rate system (1946-1972) we would expect to see greater stability in the markets. Greater stability also means lower returns. However, the mean return during this period (1.02% -- Table 3) was slightly higher than for the previous period (0.93%) which was affected by the Great Depression. This may be a reflection of greater economic activities after the Second World War and hence higher profits for companies. The mean return for December (2.46%) was the highest, followed by November (2.16%), July (1.97%) and March (1.81%). These returns are all significantly greater than zero. However, only December’s mean was significantly greater than the mean return of the other eleven months stacked, and February’s mean return was significantly lower than that of the other eleven months stacked. The higher returns for December may be a result of lower tax-loss selling at the end of the year, because of higher profits of companies. This sub-period did bring in greater stability in terms of the means of monthly returns – standard deviations ranged mostly from 3% to 4.3% for the various months. February, March and October all saw variances of returns which were significantly lower than those of the other months stacked.
TABLE 3: Descriptive Statistics and Monthly Seasonality: 1946-1972

<table>
<thead>
<tr>
<th></th>
<th>Jan</th>
<th>Feb</th>
<th>Mar</th>
<th>Apr</th>
<th>May</th>
<th>Jun</th>
<th>Jul</th>
<th>Aug</th>
<th>Sep</th>
<th>Oct</th>
<th>Nov</th>
<th>Dec</th>
</tr>
</thead>
<tbody>
<tr>
<td>count</td>
<td>324</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td>mean</td>
<td>1.02</td>
<td>1.12</td>
<td>0.10</td>
<td>0.49</td>
<td>0.34</td>
<td>0.38</td>
<td>0.97</td>
<td>0.51</td>
<td>0.24</td>
<td>0.09</td>
<td>0.16</td>
<td>0.46</td>
</tr>
<tr>
<td>median</td>
<td>1.29</td>
<td>1.94</td>
<td>0.49</td>
<td>0.65</td>
<td>0.92</td>
<td>0.61</td>
<td>0.24</td>
<td>0.34</td>
<td>0.78</td>
<td>0.75</td>
<td>0.78</td>
<td></td>
</tr>
<tr>
<td>std dev</td>
<td>3.69</td>
<td>4.28</td>
<td>2.82</td>
<td>2.75</td>
<td>3.96</td>
<td>3.95</td>
<td>3.96</td>
<td>3.95</td>
<td>3.96</td>
<td>3.95</td>
<td>3.96</td>
<td>3.96</td>
</tr>
<tr>
<td>p-value (m=0)</td>
<td>0.00</td>
<td>0.18</td>
<td>0.85</td>
<td>0.00</td>
<td>0.06</td>
<td>0.66</td>
<td>0.01</td>
<td>0.48</td>
<td>0.77</td>
<td>0.05</td>
<td>0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>p-value (t test)</td>
<td>0.90</td>
<td>0.04</td>
<td>0.14</td>
<td>0.52</td>
<td>0.36</td>
<td>0.05</td>
<td>0.19</td>
<td>0.46</td>
<td>0.11</td>
<td>0.90</td>
<td>0.14</td>
<td>0.01</td>
</tr>
<tr>
<td>p-value (F test)</td>
<td>1.6</td>
<td>0.04</td>
<td>0.03</td>
<td>0.33</td>
<td>0.34</td>
<td>0.47</td>
<td>0.42</td>
<td>0.52</td>
<td>0.20</td>
<td>0.03</td>
<td>0.25</td>
<td>0.08</td>
</tr>
<tr>
<td>mean returns</td>
<td>os</td>
<td>os</td>
<td>os</td>
<td>os</td>
<td>os</td>
<td>os</td>
<td>os</td>
<td>os</td>
<td>os</td>
<td>os</td>
<td>os</td>
<td>os</td>
</tr>
<tr>
<td>Mo effect (Mean)</td>
<td>lower</td>
<td>lower</td>
<td>lower</td>
<td>lower</td>
<td>lower</td>
<td>lower</td>
<td>lower</td>
<td>lower</td>
<td>lower</td>
<td>lower</td>
<td>lower</td>
<td>lower</td>
</tr>
<tr>
<td>Mo effect (Var)</td>
<td>lower</td>
<td>lower</td>
<td>lower</td>
<td>lower</td>
<td>lower</td>
<td>lower</td>
<td>lower</td>
<td>lower</td>
<td>lower</td>
<td>lower</td>
<td>lower</td>
<td>lower</td>
</tr>
</tbody>
</table>

Note: See note under Table 3; numbers are rounded to two decimal places.

Kruskal-Wallis nonparametric test of difference in the medians of the twelve months shows significant difference in the medians (H = 22.28, p = 0.02). Mood Median test does not show a highly significant difference in the medians of the twelve months (Chi-square = 16.74, p = 0.12). So the outcome based on non-parametric tests is mixed.

Third Sub-period: 1973-1992

The Breton Woods fixed exchange rate system was dismantled in 1972. We entered a more volatile world. The mean return for the third sub-period (1.01% -- Table 4) was significant, but slightly lower than that of the previous sub-period. January’s mean return (2.36%) was the highest, followed by December’s (1.78%), November’s (1.41%) and then June’s (1.33%). The September effect that we saw for the entire data period (September’s mean return being significantly lower than that of the other eleven months stacked) is seen for this sub-period. Evidently, the negative September effect for the entire data set is a result of this sub-period. That this sub-period became more volatile in terms of monthly standard deviations is seen from the range of standard deviations of monthly returns for various months (ranging from 3% to 8%) as well as from the overall standard deviation of the returns of this sub-period (4.69% versus 3.69% for the second sub-period). June saw lower variance of returns compared to the other eleven months stacked, and October saw higher variance. October’s higher standard deviation of monthly returns is not surprising and is born by experience of traders in recent decades.

<table>
<thead>
<tr>
<th></th>
<th>j1</th>
<th>an</th>
<th>eb</th>
<th>ar</th>
<th>pr</th>
<th>ay</th>
<th>un</th>
<th>ul</th>
<th>ug</th>
<th>ep</th>
<th>ct</th>
<th>ov</th>
<th>ec</th>
</tr>
</thead>
<tbody>
<tr>
<td>count</td>
<td>40</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>mean</td>
<td>.01</td>
<td>.36</td>
<td>.94</td>
<td>.01</td>
<td>.23</td>
<td>.40</td>
<td>.33</td>
<td>.52</td>
<td>.87</td>
<td>1.18</td>
<td>.44</td>
<td>.41</td>
<td>.78</td>
</tr>
<tr>
<td>median</td>
<td>.91</td>
<td>.96</td>
<td>.80</td>
<td>.37</td>
<td>.32</td>
<td>.70</td>
<td>.63</td>
<td>.29</td>
<td>.73</td>
<td>.20</td>
<td>.31</td>
<td>.58</td>
<td>.77</td>
</tr>
<tr>
<td>std Dev</td>
<td>.69</td>
<td>.15</td>
<td>.70</td>
<td>.56</td>
<td>.58</td>
<td>.78</td>
<td>.86</td>
<td>.50</td>
<td>.69</td>
<td>.03</td>
<td>.76</td>
<td>.23</td>
<td>.55</td>
</tr>
<tr>
<td>p-value (m=0)</td>
<td>.00</td>
<td>.10</td>
<td>.27</td>
<td>.22</td>
<td>.14</td>
<td>.11</td>
<td>.05</td>
<td>.61</td>
<td>.50</td>
<td>.21</td>
<td>.80</td>
<td>.24</td>
<td>.04</td>
</tr>
<tr>
<td>p-value (t test)</td>
<td>.31</td>
<td>.93</td>
<td>.99</td>
<td>.78</td>
<td>.64</td>
<td>.63</td>
<td>.62</td>
<td>.91</td>
<td>.02</td>
<td>.73</td>
<td>.72</td>
<td>.34</td>
<td></td>
</tr>
<tr>
<td>p-value (F test)</td>
<td>.06</td>
<td>.09</td>
<td>.06</td>
<td>.07</td>
<td>.12</td>
<td>.00</td>
<td>.43</td>
<td>.14</td>
<td>.22</td>
<td>.00</td>
<td>.28</td>
<td>.06</td>
<td></td>
</tr>
</tbody>
</table>

Note: See note under Table 3; numbers are rounded to two decimal places.

Kruskal-Wallis nonparametric test of difference in the medians of the twelve months shows no significant difference in the medians (H = 6.60, p = 0.83). Mood Median test also shows no significant difference in the medians of the twelve months (Chi-square = 8.40, p = 0.68).

Fourth Sub-period: 1993-2013

The behavior of monthly returns of the last sub-period is of special interest to us because of its proximity to present times. The mean of monthly returns (0.82%) of this sub-period (Table 5) was impacted by the Great Recession and hence is less than that of the previous sub-period. However, the median of this sub-period (1.35%) is higher than the median of the previous sub-period (0.91%). April’s mean return (2.08%) was the highest followed by October (1.78%), March (1.68%), December (1.66%), and November (1.64%).
TABLE 5: Descriptive Statistics and Monthly Seasonality: 1993-2013

<table>
<thead>
<tr>
<th></th>
<th>Jan</th>
<th>Feb</th>
<th>Mar</th>
<th>Apr</th>
<th>May</th>
<th>Jun</th>
<th>Jul</th>
<th>Aug</th>
<th>Sep</th>
<th>Oct</th>
<th>Nov</th>
<th>Dec</th>
</tr>
</thead>
<tbody>
<tr>
<td>Count</td>
<td>252</td>
<td>211</td>
<td>221</td>
<td>212</td>
<td>211</td>
<td>211</td>
<td>212</td>
<td>211</td>
<td>211</td>
<td>212</td>
<td>211</td>
<td>212</td>
</tr>
<tr>
<td>Mean</td>
<td>0.82</td>
<td>0.65</td>
<td>0.28</td>
<td>0.08</td>
<td>0.73</td>
<td>0.14</td>
<td>0.62</td>
<td>0.47</td>
<td>0.11</td>
<td>0.78</td>
<td>0.64</td>
<td>0.66</td>
</tr>
<tr>
<td>Median</td>
<td>1.35</td>
<td>1.84</td>
<td>0.81</td>
<td>0.24</td>
<td>0.58</td>
<td>0.70</td>
<td>0.47</td>
<td>0.66</td>
<td>0.81</td>
<td>0.03</td>
<td>0.03</td>
<td>0.40</td>
</tr>
<tr>
<td>StDev</td>
<td>4.30</td>
<td>3.98</td>
<td>4.21</td>
<td>2.06</td>
<td>0.95</td>
<td>0.69</td>
<td>0.67</td>
<td>0.18</td>
<td>0.69</td>
<td>0.45</td>
<td>0.67</td>
<td>0.49</td>
</tr>
<tr>
<td>p-value (m=0)</td>
<td>0.00</td>
<td>0.46</td>
<td>0.76</td>
<td>0.07</td>
<td>0.03</td>
<td>0.38</td>
<td>0.86</td>
<td>0.50</td>
<td>0.65</td>
<td>0.93</td>
<td>0.17</td>
<td>0.11</td>
</tr>
<tr>
<td>p-value (t test)</td>
<td>0.84</td>
<td>0.72</td>
<td>0.32</td>
<td>0.14</td>
<td>0.91</td>
<td>0.23</td>
<td>0.82</td>
<td>0.20</td>
<td>0.42</td>
<td>0.42</td>
<td>0.39</td>
<td>0.19</td>
</tr>
<tr>
<td>p-value (F test)</td>
<td>0.34</td>
<td>0.49</td>
<td>0.39</td>
<td>0.33</td>
<td>0.19</td>
<td>0.19</td>
<td>0.46</td>
<td>0.31</td>
<td>0.08</td>
<td>0.05</td>
<td>0.43</td>
<td>0.01</td>
</tr>
<tr>
<td>Mean Retur n</td>
<td>os</td>
<td>os</td>
<td>os</td>
<td>os</td>
<td>os</td>
<td>os</td>
<td>os</td>
<td>os</td>
<td>os</td>
<td>os</td>
<td>os</td>
<td>os</td>
</tr>
<tr>
<td>Effect (Mean)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effect (Var)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: See note under Table 3; numbers are rounded to two decimal places.

However, the mean return of none of the months was significantly higher than the mean of the other eleven months stacked. The standard deviation of the returns this sub-period (4.30%) is lower than that of the previous sub-period (4.69%). The range of standard deviations of various months was between 3% and 6% -- more compact than it was for the previous sub-period. Only the standard deviation of the returns of December was significantly lower than that of the other eleven months stacked. No other month exhibited month effect with respect to variance.

Kruskal-Wallis nonparametric test of difference in the medians of the twelve months shows no significant difference ($H = 9.96, p = 0.53$). Mood Median test also shows no significant difference in the medians of the twelve months (Chi-square = 7.43, $p = 0.76$).

Table 6: Breakdown of Monthly Returns and Standard Deviations of Four Sub-Periods

<table>
<thead>
<tr>
<th>Period</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan 1993-Dec 1995</td>
<td>1.22%</td>
<td>2.41%</td>
</tr>
<tr>
<td>Jan 1996-Dec 2002</td>
<td>0.69%</td>
<td>5.12%</td>
</tr>
<tr>
<td>Jan 2003-Dec 2011</td>
<td>0.56%</td>
<td>4.31%</td>
</tr>
<tr>
<td>Jan 2012-Dec 2013</td>
<td>1.84%</td>
<td>2.76%</td>
</tr>
</tbody>
</table>
Comparison of Four Sub-Periods

The patterns of monthly returns are different over the four sub-periods. In three separate studies [Hamid and Dhakar (2005), and (2016)] found negative September effect for the Dow Jones Industrial Average for the periods 1896 to 2002, for 1900 to 1999, and for 1896 to 2013 (the mean return for September was negative and different from the mean of the other eleven months stacked); but the September effect was found to be largely the result of the last quarter of the last century. Here also the finding is somewhat similar: the negative September effect for S&P 500 stocks is the result of the period 1973-1992; September did not produce highly pronounced negative returns in the last sub-period (1993-2013).

Based on F-test there are highly significant differences in the variances of returns of five of the six pairs that can be formed between four sub-periods (p = 0.00); the only exception being the variance of returns of the last two sub-periods that show no significant difference (p = 0.17). The standard deviation of monthly returns of the first sub-period is the highest (8.53%), followed by the third sub-period (4.69%), then followed by the fourth sub-period (4.69%), and then the second sub-period (3.69%). So the stability brought by the end of the Second World War, and the Breton Woods system of fixed exchange rates was accompanied by lower volatility in the market for large stocks (and also for small stocks as we see in a different paper). However, we do not find significant differences in the means of the six pairs formed out of the four sub-periods based on t-test assuming unequal variances (the p values are between 0.56 and 0.90). The means of the four sub-periods are: 0.93%, 1.02%, 1.01%, and 0.82% respectively, which are rather close. Also, there is no difference in the medians of the four sub-periods either based on Mood’s Median test (Chi-square = 2.51; p value = 0.47) or based on Kruskal-Wallis test (H-Statistic = 0.175; p value = 0.97). The medians of the four sub-periods are: 1.58%, 1.29%, 0.91%, and 1.35% respectively. So though large stocks have exhibited significant differences in volatilities of five of the six pairs of the four sub-periods, the means and medians of the sub-periods have not exhibited significant differences.

REPUBLICAN VS. DEMOCRATIC PRESIDENTIAL MONTHS

The typical notion people seem to have is when a Republican is in the White House the stock market will perform better than when a Democrat is the President. Policies of Republican presidents have been perceived to be more beneficial to businesses. Such eventuality should lead stock returns to be higher during Republican administrations. We test if this was indeed the case with large stocks. Niederhoffer, Gibbs and Bullock (1970) found that stock returns were much better after a Republican was elected president than after a Democrat was elected. They took short-run perspective. Johnson, Chittenden and Jensen (1999) found that for the period 1926 to 1996, stock returns based on S&P 500 were higher during Democratic presidents than during Republican presidents.

During 1926-2013 there were eight Republican presidents and seven Democratic presidents. The 518 months which had Republican presidents had a mean monthly return of 0.62% versus 1.27% for 538 Democratic presidential
months (Tables 7 and 8). The difference between the two means is significant at a p value of 0.056 in 2-tailed test assuming unequal variances. During Republican and Democratic presidential months, the standard deviations were 5.78% and 5.16% respectively. The difference between the two standard deviations is significant at a p value of 0.01. The Democratic presidential months experienced a higher median (1.58%) compared to the Republican months (0.93%) which is significantly different based on three nonparametric tests. Under Mann-Whitney test, the W-Statistic = 263571.5; p value = 0.04; under Mood’s Median test, the Chi-square = 5.78; p value = 0.02; under Kruskal Wallis test, the H-Statistic = 4.23; p value = 0.04.

**Month Effect: Republican vs. Democratic Presidential Months**

The mean of monthly returns for the Republican presidential months (0.62%) is significantly different from zero at 1% level (Table 7). The highest mean monthly return during Republican presidential months was in December (1.53%) and is significant at 3% level. The next was in November (1.48%), but significant at 8% level.

**TABLE 7: Descriptive Statistics and Monthly Seasonality During Republican Months: 1926-2013**

<table>
<thead>
<tr>
<th></th>
<th>Jan</th>
<th>Feb</th>
<th>Mar</th>
<th>Apr</th>
<th>May</th>
<th>Jun</th>
<th>Jul</th>
<th>Aug</th>
<th>Sep</th>
<th>Oct</th>
<th>Nov</th>
<th>Dec</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.28</td>
<td>0.59</td>
<td>0.41</td>
<td>0.24</td>
<td>0.52</td>
<td>0.60</td>
<td>0.92</td>
<td>2.09</td>
<td>0.28</td>
<td>0.48</td>
<td>1.60</td>
<td>1.53</td>
</tr>
<tr>
<td>Median</td>
<td>0.93</td>
<td>0.73</td>
<td>1.11</td>
<td>0.89</td>
<td>0.78</td>
<td>0.90</td>
<td>0.97</td>
<td>0.60</td>
<td>0.48</td>
<td>0.87</td>
<td>0.32</td>
<td>0.81</td>
</tr>
<tr>
<td>Std Dev</td>
<td>5.78</td>
<td>5.14</td>
<td>4.84</td>
<td>4.02</td>
<td>3.61</td>
<td>3.90</td>
<td>3.28</td>
<td>3.56</td>
<td>4.04</td>
<td>3.97</td>
<td>3.35</td>
<td>3.35</td>
</tr>
<tr>
<td>Sample Var</td>
<td>0.33</td>
<td>0.00</td>
<td>0.23</td>
<td>0.16</td>
<td>0.26</td>
<td>0.29</td>
<td>0.24</td>
<td>0.53</td>
<td>0.57</td>
<td>0.41</td>
<td>0.49</td>
<td>0.29</td>
</tr>
</tbody>
</table>

Note: See note under Table 3; numbers are rounded to two decimal places.

The mean monthly return of September was -2.09% which is significant at 4% level. The mean return for September was also significantly lower than the
mean of returns of the other eleven months stacked (the mean of returns of none of the months was significantly higher than the mean of the other eleven months stacked). So the September effect we found for the entire data period is attributable to Republican presidential months. This finding is also consistent with similar finding of Hamid and Dhakar (2005) and (2016) with DJIA stocks. July and August exhibited the highest volatility (standard deviations of monthly returns being 7.28% and 7.56% respectively) and these were higher than the volatilities of returns of the other months. These may be attributable to quarterly earnings releases during these months – but it needs further investigation. The volatilities of returns of March and December were lower than the volatilities of the other eleven months stacked. The range of standard deviations during Republican months is about the same as it was for the entire data period.

### TABLE 8: Descriptive Statistics and Monthly Seasonality During Democratic Months: 1926-2013

<table>
<thead>
<tr>
<th>Month</th>
<th>Jan</th>
<th>Feb</th>
<th>Mar</th>
<th>Apr</th>
<th>May</th>
<th>Jun</th>
<th>Jul</th>
<th>Aug</th>
<th>Sep</th>
<th>Oct</th>
<th>Nov</th>
<th>Dec</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>1.27</td>
<td>1.62</td>
<td>0.25</td>
<td>0.59</td>
<td>0.50</td>
<td>0.75</td>
<td>0.61</td>
<td>0.02</td>
<td>0.69</td>
<td>0.65</td>
<td>0.09</td>
<td>0.51</td>
</tr>
<tr>
<td>Median</td>
<td>1.58</td>
<td>2.17</td>
<td>0.65</td>
<td>0.47</td>
<td>0.63</td>
<td>0.78</td>
<td>0.96</td>
<td>0.64</td>
<td>0.31</td>
<td>0.96</td>
<td>0.33</td>
<td>0.85</td>
</tr>
<tr>
<td>St Dev</td>
<td>0.56</td>
<td>0.26</td>
<td>0.93</td>
<td>0.77</td>
<td>0.03</td>
<td>0.54</td>
<td>0.97</td>
<td>0.65</td>
<td>0.19</td>
<td>0.76</td>
<td>0.87</td>
<td>0.69</td>
</tr>
<tr>
<td>p-value (t-test)</td>
<td>0.00</td>
<td>0.02</td>
<td>0.51</td>
<td>0.04</td>
<td>0.41</td>
<td>0.06</td>
<td>0.01</td>
<td>0.33</td>
<td>0.40</td>
<td>0.13</td>
<td>0.04</td>
<td>0.00</td>
</tr>
<tr>
<td>p-value (F-test)</td>
<td>0.57</td>
<td>0.06</td>
<td>0.42</td>
<td>0.26</td>
<td>0.55</td>
<td>0.66</td>
<td>0.29</td>
<td>0.39</td>
<td>0.41</td>
<td>0.80</td>
<td>0.73</td>
<td>0.15</td>
</tr>
</tbody>
</table>

Note: See note under Table 3; numbers are rounded to two decimal places.

The higher mean of monthly returns during Democratic presidential months (1.27%) is significantly greater than zero at 0% level (Table 8). What is interesting for the Democratic months is that the monthly means of five months is significantly greater than zero as seen from the third last row of Table 8. The mean monthly return for September is not negative, which reinforces the conclusion that the negative September effect found for the entire data set is attributable to the Republican presidential months. During the Democratic months, no month had a mean return significantly different from the mean of the other eleven months.
stacked. One month saw significantly higher volatility of returns (April) and three months saw lower volatility of returns (January, February, and December) compared to the other eleven months stacked. The Democratic months were characterized by greater range of standard deviation of monthly returns (3% to 8%) compared to the Republican months (the range in this case was 4% to 8%). However, the standard deviation of Republican months (5.78%) is higher than that of Democratic months (5.16%).

The mean of monthly returns under Republican presidencies (0.62%) is significantly less than the mean of monthly returns under Democratic presidencies (1.27%); p value based on t-test is 0.056. F-test for difference in variances of returns yield a p value of 0.01. So there is significant difference in the variances. The medians of the two presidential periods (Republican period: 0.93% and Democratic period: 1.58%) are significantly different based on three nonparametric tests (Mann-Whitney W-Statistic = 263571 significant at a p value of 0.04; Mood’s Median test Chi-square = 5.78 with p value = 0.02, and Kruskal-Wallis H-Statistic = 4.27 with p value = 0.04).

**SUMMARY AND CONCLUSION**

We explored and found monthly seasonality for the entire data period: 1926-2013, and also for four sub-periods into which we divided the data series based on structural changes in the economy. We find a significant negative September effect in the monthly returns of large stocks for the full data series which was caused by the significant negative monthly returns of the third sub-period in particular, and in general by the nonsignificant negative September returns in all the other sub-periods. The mean return of December is higher than the mean of the other eleven months stacked for a p value of 0.052. April underwent significantly higher variance compared to that of the other eleven months stacked. We also find greater volatility of monthly returns in October compared to the other months in the third sub-period (significant) as well as the fourth sub-period (not significant). The higher volatility that we see in October is a recent phenomenon (since 1973); the volatility of returns of October in the second sub-period was significantly lower than that of the other eleven months stacked, and in the first sub-period, seven months had higher volatility of returns than October. We also find significantly greater volatility of monthly returns for April; January, February, and December have significantly lower volatilities compared to those of the other eleven months stacked. The fact that month effect with respect to means and volatilities is different in different sub-periods is a testament of market efficiency as well as inefficiency. We look at the performance of large stocks during Republican presidential months and Democratic presidential months. Contrary to popular belief, large stocks did better during Democratic presidents. The September effect that we found for the entire data period is attributable to Republican presidential months.
REFERENCES


WHEN RIVALS PURCHASE: THIS MEANS WAR! OR DOES IT?

David Burns
Jennifer Hutchins
Richard Mathisen
Kennesaw State University

ABSTRACT

Consumer warfare is a phenomenon in consumer culture wherein an individual who learns that a rival or peer purchased or intends to purchase a particular good has particular response patterns depending upon many elements. This paper is based upon an existing framework of defensive strategies in consumer warfare, and, given the absence of previous studies that have empirically examined consumer warfare, the authors explore important elements in the decision-making process. Specifically, the authors investigate how consumers respond to a rival when the individual has positive or negative feelings, or affect, toward that rival or peer who purchased or plans to purchase the good. Results show that individuals with higher negative affect toward a rival indicate a greater likelihood of responding with an overt warfare strategy, while those individuals with higher positive affect toward a rival indicate a greater likelihood of responding with a subtle warfare strategy.

Keywords: consumer culture, consumer behavior, consumer warfare, identity, materialism

INTRODUCTION

The purpose of this paper is to empirically examine consumer warfare through building a model of the defensive strategies. Specific attention is placed on the feelings consumers have toward their rivals and how those feelings influence their choice of defensive strategy. Some say keeping up with the Joneses is akin to warfare (von Neumann and Morgenstern 1944). Allusions to price wars, intelligence, and other similar warfare terminology have been quite common in business practice and the business press for many years (Whysall 2001). With diminishing opportunities for growth stemming from an economic slowdown in the 1970s and 80s, and stagnation of population growth in developed counties, businesses could no longer expand and grow merely as a byproduct of a growing economy or population (Kotler and Singh 1981). Instead, businesses grew at the expense of other businesses – their competition (Malhotra 1994; Ries and Trout 1986).
Within the warfare paradigm, competitors are viewed as enemies or rivals. Strategies employed for millennia on the battlefield (see Hart 1967; von Clausewitz 1966) were then employed as strategic tools in the marketplace. Attack warfare strategies were used as means through which business could be wrestled away from competitors similar to how land could be wrestled away from other countries in battlefield warfare. Defensive warfare strategies were used to retain business in the face of present or anticipated competitive attacks from other businesses. Marketing warfare proved to be a successful competitive tool (Ries and Trout 1986).

Today’s business climate is very different from that of the 1970s and 1980s. In an environment now characterized by quickly advancing technology and increasing regulation, the business environment has changed significantly (Chungyalpa and Bora 2015). Consequently, many strategies that succeeded in previous environments, such as marketing warfare, are arguably no longer effective in today’s marketplace (“Buying the Future” 2005). Today’s business environment is depicted as powerful consumers competing with and among themselves (e.g., “Crowned at Last” 2005; Hughes 2007; Kucuk and Krishnamurthy 2007). Reflecting their new-found power and importance, a focus on strategies employed by consumers in the marketplace appears to be increasingly appropriate to effectively reach today’s and tomorrow’s consumers (“Crowned at Last” 2005; Huang, Lotz, and Bon 2014).

Burns and Warren (2008) introduced the concept of consumer warfare as a means to illustrate how consumers interact and how their interactions affect their behavior in the marketplace. Based on previous research, they propose that much of consumers’ marketplace activities today is the result of consumers attempting to develop/maintain competitive identities. Consequently, consumers’ activities in the marketplace are viewed as competitive, where consumers use products as “weapons” in their endeavors to establish/maintain superiority over others (Burns and Warren 2009). Within a consumer culture, where the physical needs of most are satisfied, identity formation/maintenance becomes the primary motivation behind most consumers’ marketplace activities and may be the most fruitful focus of research by practitioners and researchers (Fitzmaurice and Comegys 2006).

The purpose of this paper is to empirically examine consumer warfare through building a model of the defensive strategies. Specific attention is placed on the feelings consumers have toward their rivals and how those feelings influence their choice of defensive strategy.

CONSUMERS’ MOTIVATIONS IN THE MARKETPLACE

Many societies today are characterized by an all-pervasive consumer culture (Easterbrook et al. 2014; Kasser and Kanner 2004; Strasser 2003). Key to consumer culture is materialism, a quality which pervades the lives of most
individuals (Belk 2015; Larsen, Sirgy, and Wright 1999). “Materialists place possessions and their acquisition at the center of their lives” (Richins and Dawson 1992, p. 304). Materialism is not a new concept, but its pervasiveness and the degree to which individuals embrace this quality is a new concept.

In consumer cultures, most of the physical needs of individuals are met; so it is initially surprising that products possess such a high level of importance within these cultures. The level of importance, however, comes from the role that products play in consumer cultures – products do not merely satisfy physical needs, but play a central role in establishing individuals’ identities or their selves (Ferreira and Scaraboto 2016; Shaw, Newholm, and Dinkinson 2006; Thompson and Loveland 2015). “Knowingly or unknowingly, intentionally or unintentionally, we regard possessions as ourselves… That we are what we have is perhaps the most basic and powerful fact of consumer behavior” (Belk 1988, p. 139). The importance of consumption activities has increased to such an extent that some suggest that individuals’ selves are formed primarily through presentation made possible through products acquired and displayed (Zukin and Maguire 2004). Ferguson suggests that

“All consumption becomes conceivable as the desire for, as well as the desire of, the self. ... We then seek, in consuming such objects (products), to incorporate an idealized self, to make the self more real, and to end the inner despair of not having a self” (Ferguson 1992, p. 27, 28).

Central to self-construction is establishing a point of difference from others (Snyder and Fromkin 1980; Tian, Bearden, and Hunter 2001) or, in other words, developing a self that is unique and, in some way, superior to that possessed by others (Schaefers 2014). Given the role of products in the process of self-construction, the acquisition of products, then, becomes the basis for interpersonal competition (Paterson 2006; Robbins 2000). This reality is not new. Over forty years ago, Fromm (1976) suggested that the acquisition of products for display can be a means by which individuals attempt to display superiority over others. Possessions, therefore, become the basis for engaging, competing with, and, at times, attempting to defeat others’ selves. The term, consumer warfare, was developed to describe this behavior (Burns and Warren 2008). Consumer warfare consists of applying the well-proven strategies of the battlefield and marketing warfare to the process of self-construction.

Rivals can take various forms – individuals with whom one directly interacts (e.g., neighbors, colleagues), groups, famous celebrities, or images created through marketing activities (Burns 2010). Given their close proximity, however, arguably, the individuals with whom one directly and frequently interacts tend to play an important role. This group of individuals will be discussed further.
REFERENT OTHERS

Referent others are individuals or groups to whom an individual looks to establish a frame of reference from which to make evaluations (Schulz 2015). Most importantly, individuals compare themselves to referent others to assess their relative personal state. The results of this comparison, in turn, drives many consumption choices (Escalas and Bettman, 2005; Tanner et al. 2008). There are several types of reference groups. One’s associative reference group, for instance, comprises those individuals who are viewed as an individual’s equals or near-equals, whereas one’s aspirational groups include those who one would most like to be. Within consumer warfare, members of one’s associative reference group most likely include an individual’s direct rivals – those individuals whose choices and perceived choices most likely prompt mimetic desire (Kaufman et al. 2012). The choices and perceived choices of members of one’s aspirational reference group, however, provide pictures of the choices that can be made to raise oneself relative to individuals in one’s associative reference group (Mead et al. 2011). When a rival makes a purchase or is perceived as having interest in a particular product, desire develops, prompting a defensive response usually consisting of attempting to copy, or preferably, surpassing the rival. In consumer cultures, this response ordinarily takes place in the marketplace (Wharff 2007). This is the essence of consumer warfare.

CONSUMER WARFARE STRATEGIES

Research on battlefield warfare provides the basis for the strategies available in consumer warfare. The strategies employed in consumer warfare are the same as those employed in battlefield warfare and in marketing warfare, but they are implemented differently and by different parties (by consumers on other consumers) (Burns and Warren 2008). The strategies are of two types – attack strategies and defensive strategies. The focus of this study is on the defensive strategies employed by consumers.

Individuals use defensive strategies when they perceive that they have been or will be attacked by another, and those strategies range from subtle to overt. Specifically, an attack is perceived with the acquisition or the anticipated acquisition of a product by a rival that is perceived to be superior in some way (e.g., a new automobile, a new house, new apparel). It should be noted that the product acquired must not necessarily be physically superior. Most often the superiority only exists aesthetically, where the “superiority” observed is only a manifestation of the desire prompted by seeing one’s rival with the product. Instead, the mere act of acquiring a product by a rival will cause the product to be viewed as superior, and the act is viewed as an attack by the rival on the individual’s self. Consequently, one’s own self can then be viewed as being adversely affected by the purchase. It should be noted that whether the purchase was intended as an attack is immaterial.
Burns and Warren (2008) identify six consumer warfare defense strategies, each based on a warfare/marketing warfare strategy. Two subtle strategies (Satisfaction and Check-Out) involve not responding to the attack. With the Satisfaction defense, the attack is not viewed as threatening the relative status of one’s self—hence, no response is required. With the Check-Out defense, the attack is viewed as one where there is not an adequate response, so there is no need to take any action, but to surrender. Four overt strategies include: first, the No-You-Don’t strategy which involves an anticipatory attack—an attack from a rival is anticipated, so the individual responds by attacking first. Second, and very similar, the I’ll Get You strategy, involves a retaliatory attack—an attack is perceived, so the individual responds by attacking back. If a rival makes a purchase, a similar or a superior purchase is made in an attempt to negate the effect of the rival’s purchase. The third and fourth strategies focus on alternative product categories as bases for self-construction as a way to blunt future attacks involving products from those categories (New Reference) or to prevent future attacks from a new product area (Innovation).

**CHOICE OF DEFENSIVE STRATEGY**

Consumer warfare is inherently an interpersonal activity, and one’s use of a defensive strategy is no exception. DeRidder, Schrijver, and Rijsman (1999) observed that the recipient of an attack will respond based on their interpretation of the social context in which the attack was made. Given that a defensive strategy is directed toward an individual viewed as a rival, the nature of the affect held by the individual toward their rival can be expected to have an effect on the specific defensive strategy employed. Previous research examining the reactions of individuals to attacks seems to suggest this connection. Several (e.g., Crossley 2008; Juujärvi et al. 2001; Orbach 1978), for instance, observed that the qualities of attackers as perceived by their victims appear to affect the nature of their reactive responses.

Affect has been shown to be comprised of two independent, non-related factors—positive affect and negative affect (Watson and Naragon 2009). Regarding positive affect, Nesdale, Rule, and Hill (1978) observed that victims tend to be less likely to ascribe perceived attacks to an attacker’s personality for attackers for whom they possess positive affect. Consequently, the presence of positive affect toward one’s attacker can be expected to blunt the nature of one’s response to an anticipated or perceived attack. In other words, the victim will be less likely to respond to an attack with an overt defense. Hence, individuals who perceive that they are the target of a consumer warfare attack can be expected to be less likely to respond with overt consumer warfare defensive strategies if they possess positive affect toward their perceived attacker.

Negative affect toward an attacker can also be expected to affect one’s response to a perceived attack. Crossley (2008), for instance, noted that individuals
tend to react more overtly when they possess negative affect toward their attackers. Hence, individuals who perceive that they are the target of a consumer warfare attack can be expected to be more likely to respond with overt consumer warfare defensive strategies if they possess negative affect toward their perceived attacker.

Our focus is on examining individuals’ choice of defensive strategy in consumer warfare. When an individual perceives a threat from a rival, the choice of defensive strategy to employ can logically be expected to be affected by the nature of the relationship that exists between the individual and their rival. The purpose of this paper is to begin to empirically examine consumer warfare defensive strategies, with specific attention on the effects of consumers’ affect toward their rival. In the situation of a perceived attack by a rival, how do individuals choose their response? The effects that one’s positive and negative affect toward their rival has on their attitudes towards defensive strategy alternatives are examined.

THE STUDY

The nature of the relationship between individuals and their perceived attackers can logically be expected to have an effect on the response, or the choice of defense strategy to be pursued. Specifically, it is logical to expect that the response to a perceived attack originating from an individual with whom one has more positive affect will be less overt than when the perceived attack originates from an individual with whom one has lesser positive affect. Similarly, it is logical to expect that the response to a perceived attack originating from an individual with whom one has more negative affect will be more overt than when the perceived attack originates from an individual with whom one has lesser negative affect.

HYPOTHESES

It is hypothesized that individuals’ choice of defensive strategy differs based on the level and types of affect they possess toward their rivals. Two of the strategies, the No-You-Don’t and I’ll Get You strategies, differ only in the timing of the response – whether the defense is employed before or after the actual attack. Given the difficulty in measuring the timing of the response, only the I’ll Get You strategy was examined in this study. Of the remaining defensive strategies, three involve an overt response. Given the above discussion, the overt strategies can logically be expected not to be viewed as viable options to employ against attackers viewed with positive affect. Consequently, it is logical to expect no or weak negative relationships between the levels of positive affect held toward one’s rival and likelihood to respond with a defensive strategy involving an overt response.

H1a: Individuals with greater positive affect toward a rival can be expected to possess little likelihood of responding to a perceived attack by their rival.
by purchasing similar or superior products than individuals with lesser positive affect toward their rival.

**H1b:** Individuals with greater positive affect toward a rival can be expected to possess little likelihood of responding to a perceived attack by their rival by purchasing products from different product categories as a basis for self-construction than individuals with lesser positive affect toward their rival.

**H1c:** Individuals with greater positive affect toward a rival can be expected to possess little likelihood of responding to a perceived attack by their rival by purchasing products from different product categories to prevent their rivals from purchasing such products to launch an attack upon themselves than individuals with lesser positive affect toward a rival.

The final two strategies involve less overt responses. Satisfaction, or overlooking the attack, can logically be expected to be a preferred option to be taken against individuals one possesses positive affect.

**H1d:** Individuals with greater positive affect toward a rival can be expected to possess a significantly greater likelihood of responding to a perceived attack by their rival by doing nothing than individuals with lesser positive affect toward a rival.

The Check-Out defense involves acquiescing to the rival or admitting defeat. This does not appear to be logically expected to be a preferred option to be taken against individuals with whom one possesses positive affect.

**H1e:** Individuals with greater positive affect toward a rival can be expected to possess little likelihood of responding to a perceived attack by their rival by admitting defeat than individuals with lesser positive affect toward a rival.

When examining the level of negative affect toward one’s rival, the preferred responses by an individual to a perceived attack can be expected to be to respond to the attack in an overt fashion. The presence of negative affect will lead individuals to respond to perceived attacks in a more direct fashion.

**H2a:** Individuals with greater negative affect toward a rival can be expected to possess a significantly greater likelihood of responding to a perceived attack by their rival by purchasing similar or superior products than individuals with lesser negative affect toward their rival.

**H2b:** Individuals with greater negative affect toward a rival can be expected to possess a significantly greater likelihood of responding to a perceived attack by their rival by purchasing products from different product categories as a basis for self-construction than individuals with lesser negative affect toward their rival.

**H2c:** Individuals with greater negative affect toward a rival can be expected to possess a significantly greater likelihood of responding to a perceived attack by their rival by purchasing products from different product categories to prevent their rivals from purchasing such products to launch
an attack upon themselves than individuals with lesser negative positive affect toward a rival.

The final two strategies involve less overt responses. Satisfaction, or overlooking the attack, can logically be expected not to be a preferred option to be taken against individuals one possesses negative affect. It is unlikely that individuals will overlook attacks from individuals with whom they possess negative affect.

**H2d:** Individuals with greater negative affect toward a rival can be expected to possess a significantly lower likelihood of responding to a perceived attack by their rival by doing nothing than individuals with lesser negative affect toward a rival.

Check-Out defense, however, logically appears to be the preferred option to be taken against individuals one possesses negative when the more overt strategies to not appear to be reasonably possible.

**H1e:** Individuals with greater negative affect toward a rival can be expected to possess significantly greater likelihood of responding to a perceived attack by their rival by admitting defeat than individuals with lesser negative affect toward a rival.

The hypothesized model is presented in Figure 1.

Figure 1
Hypothesized Model
METHODOLOGY

The questionnaire included instruments to measure the level of positive affect of study participants toward their rival, the level of negative affect of study participants toward their rival, and to measure their attitudes toward the possible consumer warfare defensive strategies in response to a perceived attack from their rivals. Given the absence of previous studies that have directly empirically examined consumer warfare, instruments had to be developed for use in the study. As previously mentioned, given the difficulty in developing items to clearly differentiate between the No-You-Don’t and I’ll-Get-You strategies, only items measuring the I’ll Get You strategy were included in the instrument. The items were developed by a group of marketing faculty members knowledgeable of consumer warfare and the accompanying strategies, and the items were informally reviewed by students for clarity. The indicator scales are shown in Figure 1. Respondents were asked to agree or disagree with a five point scale from 5 Strongly Agree to 1 Strongly Disagree.

SAMPLE

The sample was comprised of 262 students. The use of a sample comprised of college students is appropriate for this exploratory study. Individuals in this cohort are particularly influenced by the choices of others, including choices and behavior in the marketplace (Roberts, Manolis, and Tanner 2008). It also is during adolescence and early adulthood that individuals form their identities (Klimstra et al. 2010). Hence, it is logical to expect that college students may be more likely to engage in consumer warfare, especially defensive consumer warfare strategies, than other groups of individuals. Perceived attacks on the selves of college students resulting from the purchase of products by one’s referent others would seem to be viewed as particularly important, often resulting in the need for a defensive response.

Students attending several marketing classes at a university located in the Midwest were asked to complete a questionnaire including the described scales. The questionnaire was administered in classroom settings. No nonresponse was noted. Students are an appropriate audience for this study. Although there is a movement among a small number of young people against global capitalism, Hooks (2000) writes that young people today identify themselves by their possessions, such as clothes, phones, cars, and so on, instead of putting forward their personality and character. This perspective captures the essence of the consumer culture discussed herein.
FINDINGS

The research objective was to build a model of Consumer Warfare Theory. No extant theories have been operationalized in the literature. Hair et al. (2011; p. 144) recommend that “…If the research is exploratory or an extension of an existing structural theory, select PLS-SEM.” SmartPLS (Ringle et al 2015) was used to analyze the hypothesized model. Hair et al. (2014) suggest that the PLS model be initially evaluated for the indicator outer loadings. To be included these should be above a threshold value of 0.708. Next the composite reliability, convergent validity and discriminant validity were assessed. Convergent validity is measured by the Average Variance Extracted (AVE). Discriminant validity has been traditionally measured by the Fornell Larker Criteria (FLC). Henseler, Ringle and Sarstedt (2015) suggest that the Heterotrait-Monotrait (HTMT) ratio further identifies the lack of discriminant validity and is a superior approach to the Fornell Larker criteria. Voorhees et al. (2015) tested four discriminant validity measures for marketing and concluded that the HTMT method “offers the best balance between high detection and low arbitrary violation (p. 119).” HTMT measures the correlations of the indicators across constructs. If the HTMT is above 0.90 (HTMT_{0.90}) then combining constructs should be considered (Henseler, Ringle and Sarstedt 2015). If the resulting model has all HTMT ratios of less than 0.90, the model has discriminant validity. Finally, Henseler and Sarstedt (2013) suggest that goodness of fit is not appropriate for assessing whether a model is adequate using PLS-SEM.

MODEL ANALYSIS

One indicator (I feel much better by purchasing the same product for myself.) in the hypothesized model was eliminated in the final model. The indicator had an outer loading of 0.622 which was less than the threshold value of 0.708. After the outer loadings were evaluated, the HTMT ratio indicated a discriminant validity issue with the New Reference and Innovation constructs where the ratio was 0.9091. At the HTMT_{0.90} and HTMT_{0.85} levels these constructs appeared not to have discriminant validity. In the final model these constructs were combined, as suggested by Henseler, et al. (2015), into a single construct. All the indicators except the prior elimination of Q4 were retained the model.

The outer loadings of the indicators in the final model are illustrated on the arrows from the indicators to the constructs in Figure 2. All indicators loaded higher than 0.708 on their respective construct.

The model path coefficients are shown on the arrows in Figure 2. The largest path coefficient from Positive Affect was to the Satisfaction construct. The largest path coefficient from Negative Affect was to the Checkout construct with I’ll Get You and New Reference/Innovation as the second and third largest Bootstrapping with 5000 repeated samples was used to develop the sampling
distribution (Hair et al. 2014). The paths from Positive Affect to I’ll Get You (p = 0.3178) and Checkout (p = 0.4400) are not statistically significant at the 0.05 level. The nonsignificant paths are not shown in Figure 2.

Figure 2
Final Model

The model has both Construct Reliability and Validity. As shown in Table 2, the Cronbach’s Alpha for the constructs is well above 0.60. The average variance extracted was 0.60 and higher. The composite reliability for the constructs is above 0.80.

The model has discriminant validity as shown in Table 3. The HTMT0.90 ratios confirm that the model exhibits discriminant validity.
Table 2. HeteroTrait-MonoTrait Ratio

<table>
<thead>
<tr>
<th>Satisfaction</th>
<th>I'll Get You</th>
<th>New Ref Innovation</th>
<th>Checkout</th>
<th>Pos Affect</th>
<th>Neg Affect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Satisfaction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I'll Get You</td>
<td>0.4868</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New Reference/Innovation</td>
<td>0.5700</td>
<td>0.7312</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Checkout</td>
<td>0.1458</td>
<td>0.4529</td>
<td>0.3036</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive Affect</td>
<td>0.4100</td>
<td>0.1853</td>
<td>0.2919</td>
<td>0.1119</td>
<td></td>
</tr>
<tr>
<td>Negative Affect</td>
<td>0.3450</td>
<td>0.5207</td>
<td>0.4314</td>
<td>0.6301</td>
<td>0.3011</td>
</tr>
</tbody>
</table>

Table 3 presents the construct coefficients of determination ($R^2$). Bootstrapping with 5000 samples was used to estimate the significance of the $R^2$ values. The construct $R^2$ values are all significant at the 0.05. The coefficients of determination are illustrated in the construct circles in Figure 2.

<table>
<thead>
<tr>
<th>Table 3. Construct R^2 Bootstrapping Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original Sample $R^2$</td>
</tr>
<tr>
<td>------------------------</td>
</tr>
<tr>
<td>Satisfaction</td>
</tr>
<tr>
<td>I'll Get You</td>
</tr>
<tr>
<td>New Reference/Innovation</td>
</tr>
<tr>
<td>Checkout</td>
</tr>
</tbody>
</table>

Note: All $p$-values are significant at a level of .01 or higher

**DISCUSSION**

Humans are inherently social beings (Young 2008). Consequently, interactions among individuals take priority positions in individuals’ lives and influence many of their choices and actions. The outcome of that is that for many purchases, the physicality, or the physical benefits, of a physical product is virtually meaningless, but that the meaning ascribed to products often take precedence. This reality has been well-accepted in marketing for some time. What has not received research attention is the specifics of how other individuals can affect the preferences of individuals to pursue various strategies within the marketplace. Specifically, how do consumers respond when the actions of another (one’s rival) are perceived to affect one’s self?

This study considered the effect that positive affect and negative affect toward one’s perceived attacker (their rival) may have on an individual’s preference for reaction. After eliminating one strategy and combining another to address discriminant validity issues, each of the proposed hypotheses was
supported. Specifically, individuals who held positive affect toward their rival preferred the strategy of Satisfaction (do nothing) when faced with a perceived attack. Individuals who held negative affect toward their rival, however, were found to prefer to pursue the more overt defense strategies of I’ll Get You or New Reference/Innovation. Also, they were observed not to prefer the Satisfaction defense, but were observed to prefer the Check-Out strategy.

The findings suggest that the marketplace choices of individuals are affected by the choices of others – their rivals. Furthermore, the findings suggest that the affect one has toward a rival will affect how they will react. Moreover, the findings are consistent with interpersonal competition and provide support for the defense strategies in consumer warfare as proposed by Burns and Warren (2008).

The results appear to be important to marketers. Consumer warfare is viewed by some as a major driver of retail sales in a consumer culture (Burns and Warren 2008). Hence, anything that would hinder consumers’ direct overt responses to perceived attacks from their rivals may potentially limit their involvement in the marketplace. In other words, any such hindrances may result in lesser purchasing activity. The results of this study suggest that positive affect towards one’s rivals may be such a hindrance. Consequently, marketers may possess a motivation to attempt to lessen the positive affect consumers may possess toward their rivals. The results of the study also suggest that increasing negative affect toward one’s rivals may also be instrumental in encouraging a market-based reaction from individuals who perceive that they have been attacked. When one views some of the existing promotion activity by some marketers, this approach seems to be imbedded in their strategy.

If corroborated by further research, the findings seem to suggest that high positive affect toward one’s rivals may tend to thwart overt responses to perceived attacks by one’s rivals. If the success of a business depends on the use of consumer warfare defensive strategies (products are purchased by individuals as a means of responding to the purchases of a rival), the existence of lower positive affect toward rivals may be advantageous since it may result in higher sales. Furthermore, the presence of negative affect toward rivals appears to increase the likelihood of an overt marketplace reaction involving the acquisition of additional product(s).

Ignoring the obvious ethical ramifications, when viewing sales revenue, it may be advantageous for some businesses to try to minimize positive affect and maximize negative affect between an individual and their rival to foster additional sales opportunities. It appears that some consumer goods companies have used this very approach. The Pantene “Don’t hate me because I’m beautiful” campaign may be a classic example. The message “Don’t hate me” is structured to lessen positive affect and to foster negative affect toward the individual in the ad (who is depicting one’s rival) to prompt an overt response consisting of purchasing Pantene shampoo. Similarly, in Samsung’s campaign for the introduction of the Galaxy S2
smart phone, individuals waiting in a long line for the new iPhone express opinions/actions ranging from indifference to mockery and disgust. It was done in a fashion to ridicule the individuals waiting in line for an iPhone in a successful attempt to depict iPhone owners as individuals to whom positive affect would be misplaced and depicting the Samsung phone as a successful “weapon” to employ as a part of an overt attack against rivals who own iPhones. Similarly, an older campaign by Wilcom Cellular depicted rivals uppity snobs (low affect) to whom one needed to make a purchasing statement (which in their case was obviously the purchase of the newest cellphone).

LIMITATIONS

Several factors limit the generalizability of the findings from this study. First, as an initial exploratory study, the instruments used have not been subject to extensive validity testing. Second, the sample was constrained demographically. Although necessary to control for demographic differences, the generalizability of the findings to other populations is unknown. Third, the study examined attitudes toward the defensive consumer warfare strategies and did not examine behavioral responses. The relationship between attitudes toward the strategies and actual choice of strategy is unknown. Nonetheless, these findings begin to reveal the relationship between feelings toward one’s rival and feelings toward consumer warfare strategies.

REFERENCES

Burns, Hutchins and Mathisen


Burns, Hutchins and Mathisen

Validation,” *Journal of Consumer Research*, 19 (December), 303-316.


DOES UNWINDING CARRY TRADE LEAD TO PROFITABLE REVERSE CARRY TRADE? THE CASE OF YEN

Minje Jung
University of Central Oklahoma

ABSTRACT: The returns on yen carry trade during the period October 2007 through August 2011 were found to be negative defying the conventional notion of highly speculative but profitable short-term yen carry trade. Carry trade is a bet against the Interest Rate Parity under the interest rate differentials and temporary mispricing of the target currency. 2008 Global financial crisis narrowed the interest rate differential between Japan and the U.S., and yen carry trade unwound during the period. The objective of the paper is to investigate whether the unwinding yen carry trade was profitable for the U.S. reverse yen carry traders who would have borrowed dollar before conversion into yen and investment in yen that appreciated despite low interest rate. I found that the yen carry trades were not profitable with negative returns for 1-month, 3-month, and 1-year investment horizons. I also investigated whether interest rate differential and market volatility as measured in VIX can explain the positive returns on 1-month, 3-month, and 1-year horizons. Volatility index (VIX) is found to be negatively related to the yen reverse carry trade returns. As market becomes more volatile, the return on the reverse yen carry trade declines.

Key Words: Yen Carry Trade, Reverse Carry Trade, Uncovered Interest Parity (UIP), London Interbank Offered Rate (LIBOR)

INTRODUCTION

The objective of the paper is to investigate whether the reverse yen carry trade would result in positive returns over one-month, three-month, and one-year investment horizons, and how they are related to two explanatory variables: VIX and interest rate differences. The reverse yen carry trade is the exact inversion of process of the yen carry trade. Interest rate in Japan has been consistently lower than in the U.S., and the interest rate differential between the two has been prompting a common short-term speculation called yen carry trade that involves borrowing in yen in at a low interest rate, selling the yen for dollar to invest in the U.S. higher-yielding assets. Because UIP predicts that the currency with the higher interest rate will depreciate against the currency with the lower rate, the carry trader is actually betting against UIP (Colavecchio, 2008). The investor in yen carry
trade seeks steady and small carry trade returns from the interest rate differential betting that dollar (i.e., the higher-yielding target currency) would appreciate rather than depreciate, while yen would depreciate rather than appreciate due to a large volume of carry trades that sell yen to buy dollar. The yen carry trade relies on relative stability of the higher U.S. short-term interest rates and a favorable exchange rate movement for dollar. As long as the dollar appreciates against the yen, the yen carry trade would result in profit because the investment proceeds in dollar would be converted into more yen.

The danger of the yen carry trade is that a large sudden loss is possible when the dollar depreciates and the yen appreciates consistent with the prediction of the UIP. To the extent that the yen carry trade investors would suffer loss (and unwind investment positions) as the yen appreciates enough to offset the interest differential earned on the dollar investment, investors who reversed the entire process should enjoy profit by borrowing dollar before converting into yen to invest in yen that appreciated despite low interest rate. Reverse yen carry trade is economically sensible to the extent that the rate of appreciation in yen exceeds the interest rate differential between dollar and yen.

According to Colavecchio (2008), carry trades tend to be pursued only when the interest differential is wide enough to compensate for the foreign exchange risk being taken, and she confirms that an increase in exchange rate volatility and a decrease in expected interest rate differentials lead to a higher probability of carry trade unwinding. She also found that in periods of high volatility of exchange rates, the risk to carry trade increases.

The CBOE Volatility Index® (VIX®) is a key measure of investor sentiment and market expectations of near-term volatility conveyed by S&P 500 stock index option prices. When VIX increases, it is less attractive to engage in reverse carry trade, and hence reverse carry traders will require higher expected returns on it. Jung and Black (2016) found that the yen carry trade was a money-losing game after the 2008 global financial crisis due to a narrower interest rate differential between Japan and the U.S. and higher exchange rate risk driven by higher market volatility.

**LITERATURE REVIEW**

Fama (1984) found an inconsistency prevalent in short-term foreign exchange market that currencies with high yield tend to appreciate and the low yield currencies depreciate: the opposite to the prediction of UIP, and he called it “forward premium puzzle”. Despite many academic research attempts to explain the puzzle, it still remains unresolved, thus
spurring short-term investors to exploit arbitrage profit opportunities by attempting currency carry trade.

Flood and Rose (2002) tested for uncovered interest parity (UIP) using daily data for 23 developing and developed countries during the crisis-strewn 1990s. Their study found that UIP works better on average in the 1990s than in previous and UIP systematically failed for fixed and flexible exchange rate countries than for crisis countries implying that high yield currencies appreciate while those low yield currencies appreciate as opposed to the prediction of the UIP.

Shapiro (2010) stated that the carry trade is one example of capital movements attempted to take advantage of interest rate differentials that drives exchange rates in the opposite direction to that predicted by the UIP. He further stated that the existence of the carry trade also indicates that many investors believe that they can profitably arbitrage interest rate differentials across countries on an unhedged (or uncovered) basis.

Colavecchio (2008) viewed the carry trade as a bet against UIP underpinned by low exchange rate volatility and interest rate differentials. Specifying a Markov-switching version of the UIP equation with time-varying transition probabilities, she found the presence of carry trade regime whose features are compatible with carry trade activities. She investigated the role that exchange rate volatility plays in the build-up of yen carry trade positions, and she found that both an increase in exchange rate volatility and a decrease in expected interest rate differentials lead to carry trade unwinding.

Mark Dow (2013) argued that ex-ante return on the carry trade mainly comes from the depreciation/appreciation of the pair of currencies driven by increased volatility of the foreign exchange rate rather than interest rate differentials because overall interest rates across countries are too low to produce a significant carry (i.e., interest rate differential) that plays a major role in carry trade. Despite its lack of thorough empirical testing, his argument points to an important tenet that increased price volatility causes the foreign exchange markets to deviate from the UIP condition.

In order to track carry trade and to shed light on its impact on exchange rate changes, Beranger et al. (1999) explored how the yen carry trade affected exchange rate volatility in 1998. Their study showed that the profitability of the yen carry trade and the daily changes of exchange rate are highly correlated and the former is the major determinant of the volatility of the latter.

Nishigaki (2007) examined the relationship between the yen carry trade and the financial factors in the U.S. and the Japan, and identified two
factors that affect the yen carry trade: the change in yield differentials between the two countries and the change in investors’ risk aversion. His study further examined the consequences of carry trade unwinding concluding that: (1) the yield difference between the U.S. and Japan is so wide that a yield differential shock would not affect the yen carry trade in any significant way, (2) the U.S. stock market has a dominant impact on the yen carry trade, and (3) the position taken by investors of the speculative yen carry trade affect the exchange rate such that an increase (decrease) in the long position of the yen cause the dollar to depreciate (appreciate) against the yen.

Cairns, Ho, and McCauley (2007) found that low-yielding currencies tend to appreciate and high-yielding currencies tend to depreciate during times of heightened global equity and bond market volatility; hence, carry trades that target high-yielding currencies and use low-yielding currencies are susceptible to financial uncertainty.

Lustig and Verdelhan (2007) developed a consumption-based model and showed that aggregate consumption growth risk explains why low interest rate currencies do not appreciate (and why high interest rate currencies do not depreciate) as much as the interest rate differential. It is shown that the high interest rate currencies depreciate on average when domestic consumption growth is low and the low interest currencies appreciate under the same conditions. Thus, low interest rate currencies provide a hedge against consumption growth risk.

Lustig, Roussanov, and Verdelhan (2011) also showed that global equity market volatility is related to the variation in currency returns between high and low interest rate currencies. It was believed that the level of risk aversion of carry traders is directly related to financial volatility and the risk aversion of investors in other financial arenas.

Gagnon and Chaboud (2007) suggested that carry trades are subject to contagion to the downside. When exchange rates move counter to expectations of carry traders, lenders will require additional collateral. To meet the margin calls, some carry traders are forced to unwind their positions. This unwinding will lead to further appreciation of the funding currency and depreciation of the target currency, leading to further unwinding. The process becomes a chain reaction similar to a flu epidemic.

Using a regime-changing model, Ichiue and Koyama (2008) examined how exchange rate volatilities cause UIP to fail. Their findings confirm market participants’ view that in a low-volatility environment, unwinding of the short-term carry trades significantly impact exchange rate returns.
As for the investment time horizon, Madura (2012) viewed carry trade as a short-term investment executed for only a day or for several months, Chinn and Meredith (2004) suggested that only short-term carry trade is feasible as UIP has been almost universally rejected in studies of exchange rate movements that used short horizon data. Using interest rates on longer-maturity bonds, they found much more support for uncovered interest parity implying that long-term carry trade is not feasible.

**METHODOLOGY**

Two key parameters of UIP, exchange rate and interest rate difference between Japan and the U.S., appear to have changed significantly. As reported in Table 1, yen appreciated (i.e. dollar depreciated) significantly after the 2008 global financial crisis from ¥107.8485/$ to ¥89.2825/$ with an increase in volatility from the standard deviation 4.3013 to 6.6011. Table 2 clearly indicates that interest rate differences between Japan and the U.S. as measured in LIBOR, narrowed substantially for 1-month, 3-month, and 1-year. For instance, LIBOR 1-year interest rate differential between Japan and the U.S. shrank from -2.40907% to -0.5891%. Both shifts in the parameters, i.e. yen appreciation with higher volatility and narrower interest differentials set an ominous stage for the yen carry traders as yen carry trade would unwind. This unwinding carry trade signifies a potential short-term speculative profit making through reverse carry trade that involves borrowing in a currency that is expected to depreciate (despite higher interest rate) to invest in a currency that is expected to appreciate (despite lower interest rate).
Table 1
Yen Spot Exchange Rate

<table>
<thead>
<tr>
<th></th>
<th>Before crisis</th>
<th>after crisis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>¥107.8485/$</td>
<td>¥89.2825/$</td>
</tr>
<tr>
<td>S.D.</td>
<td>4.3913</td>
<td>6.6011</td>
</tr>
<tr>
<td># of weeks</td>
<td>46</td>
<td>146</td>
</tr>
</tbody>
</table>

Table 2
Interest Rate Difference \(i_y-i_s\) (%)

<table>
<thead>
<tr>
<th></th>
<th>1-month</th>
<th>3-month</th>
<th>1-year</th>
</tr>
</thead>
<tbody>
<tr>
<td>before crisis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>-2.70193</td>
<td>-2.61909</td>
<td>-2.40907</td>
</tr>
<tr>
<td>S.D.</td>
<td>1.026846</td>
<td>0.996167</td>
<td>0.780874</td>
</tr>
<tr>
<td># of weeks</td>
<td>46</td>
<td>46</td>
<td>46</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>1-month</th>
<th>3-month</th>
<th>1-year</th>
</tr>
</thead>
<tbody>
<tr>
<td>after crisis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>-0.25571</td>
<td>-0.34078</td>
<td>-0.5891</td>
</tr>
<tr>
<td>S.D.</td>
<td>0.602198</td>
<td>0.663651</td>
<td>0.592031</td>
</tr>
<tr>
<td># of weeks</td>
<td>146</td>
<td>146</td>
<td>146</td>
</tr>
</tbody>
</table>

Variables and Regression Model: Reverse yen carry trade returns are calculated every Monday for 1-month, 3-month, and 1-year speculation based upon yen spot exchange rates and LIBOR interest rate differentials between Japan and the U.S. from October 1, 2007 and August 15, 2011. Money market interest rates and foreign exchange rates were collected from the Financial Times markets data available at http://www.ft.com/home.us. Volatility Index (VIX) was collected every Monday from the CBOE website for the period of this study. VIX is a measure of changing investor sentiment and market expectations market volatility. When VIX increases,
it is less favorable to launch reverse carry trades, and hence reverse carry traders will end up with lower-than expected returns on it.

The dependent variables $RET_{1,t}, RET_{3,t}$ and $RET_{12,t}$ are 1-month, 3-month, and 1-year (i.e., 12-month) returns, respectively on the reverse yen carry trade, and they are defined as:

$$RET_{1} = \ln(S_{t}) - \ln(S_{t+1\text{month}})$$

$$RET_{3} = \ln(S_{t}) - \ln(S_{t+3\text{month}})$$

$$RET_{12} = \ln(S_{t}) - \ln(S_{t+1\text{year}})$$

where $\ln$ is natural logarithm syntax. $S_{t}$ is spot exchange rate of the yen at the time of carry trade, and $S_{t+1\text{month}}, S_{t+3\text{month}},$ and $S_{t+1\text{year}}$ are realized spot exchange rates after 1-month, 3-month, and 1-year, respectively.

The independent variables are interest rate differential and VIX. $i_{¥} - i_{$}$ is interest differential between Japanese and the U.S. 1-year LIBOR interest rates. The prediction of the Interest Rate Parity is supported by the findings of Caim et al (2007) that: low-yielding currencies tend to appreciate and high-yielding currencies tend to depreciate during times of high volatility in global financial and foreign exchange markets. Their findings are consistent with my conjecture that yen carry trade is likely to unwind, thus reverse process of the transaction would result in short term positive return.

The following regression model that is typically used for empirical testing of the Interest Rate Parity, is employed for regression analysis for this study.

$$RET_{n,t} = \alpha + \beta_{1} \cdot (i_{¥} - i_{$}) + \beta_{2} \cdot \text{VIX}_{t} + \xi_{t}.$$  

For the Interest Rate Parity to hold, coefficient $\beta_{1}$ should be a statistically significantly positive value close to 1, and $\alpha = 0$. If the coefficient is zero or negative, the Interest Rate Parity will not hold and reverse carry trades will be profitable. $\beta_{2}$ the coefficient for VIX is expected to be positive.
EMPIRICAL RESULTS

Table 4
Descriptive Statistics of Reverse Yen Carry Trade Returns

<table>
<thead>
<tr>
<th></th>
<th>1-month</th>
<th>3-month</th>
<th>1-year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.00861</td>
<td>0.03717</td>
<td>0.099344</td>
</tr>
<tr>
<td>S.D.</td>
<td>0.00251</td>
<td>0.00481</td>
<td>0.004443</td>
</tr>
<tr>
<td>Min</td>
<td>-0.09561</td>
<td>-0.12926</td>
<td>-0.034690</td>
</tr>
<tr>
<td>Max</td>
<td>0.12498</td>
<td>0.21893</td>
<td>0.238649</td>
</tr>
</tbody>
</table>

As reported in Table 4, the average returns on the reverse yen carry trade are consistently positive during the sample period for 1-month, 3-month, and 1-year investment horizons. This is an indication that the reverse yen carry trades were profitable, yielding on average positive returns during the sample period.

Table 5
Regression Results
1-month Reverse Carry Trade

Regression Statistics

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple R</td>
<td>0.165991</td>
</tr>
<tr>
<td>R Square</td>
<td>0.027553</td>
</tr>
<tr>
<td>Adjusted R Square</td>
<td>0.017097</td>
</tr>
<tr>
<td>Standard Error</td>
<td>0.034147</td>
</tr>
<tr>
<td>Observations</td>
<td>189</td>
</tr>
</tbody>
</table>

ANOVA

<table>
<thead>
<tr>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>Significance F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>2</td>
<td>0.006145</td>
<td>0.003073</td>
<td>2.635041</td>
</tr>
<tr>
<td>Residual</td>
<td>186</td>
<td>0.216883</td>
<td>0.001166</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>188</td>
<td>0.223029</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Coefficients

<table>
<thead>
<tr>
<th></th>
<th>t Stat</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.003491</td>
<td>0.544944</td>
</tr>
<tr>
<td>¥ - i$</td>
<td>-0.45138</td>
<td>-2.27114</td>
</tr>
<tr>
<td>VIX</td>
<td>0</td>
<td>0.231627</td>
</tr>
</tbody>
</table>
### 3-month Reverse Carry Trade

**Regression Statistics**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple R</td>
<td>0.369482</td>
</tr>
<tr>
<td>R Square</td>
<td>0.136517</td>
</tr>
<tr>
<td>Adjusted R</td>
<td>0.126358</td>
</tr>
<tr>
<td>Standard Error</td>
<td>0.06129</td>
</tr>
<tr>
<td>Observations</td>
<td>173</td>
</tr>
</tbody>
</table>

**ANOVA**

<table>
<thead>
<tr>
<th></th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>Significance F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>2</td>
<td>0.09414</td>
<td>0.04707</td>
<td>13.43849</td>
<td>0</td>
</tr>
<tr>
<td>Residual</td>
<td>170</td>
<td>0.595444</td>
<td>0.0033503</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>172</td>
<td>0.689584</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Coefficients**

<table>
<thead>
<tr>
<th></th>
<th>t Stat</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>5.693517</td>
<td>0</td>
</tr>
<tr>
<td>i - i$</td>
<td>-3.79683</td>
<td>0</td>
</tr>
<tr>
<td>VIX</td>
<td>-3.99603</td>
<td>0</td>
</tr>
</tbody>
</table>

### 1-year Reverse Carry Trade

**Regression Statistics**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple R</td>
<td>0.653306</td>
</tr>
<tr>
<td>R Square</td>
<td>0.426809</td>
</tr>
<tr>
<td>Adjusted R</td>
<td>0.418502</td>
</tr>
<tr>
<td>Standard Error</td>
<td>0.040229</td>
</tr>
<tr>
<td>Observations</td>
<td>141</td>
</tr>
</tbody>
</table>

**ANOVA**

<table>
<thead>
<tr>
<th></th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>Significance F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>2</td>
<td>0.166298</td>
<td>0.083149</td>
<td>51.3787</td>
<td>0</td>
</tr>
<tr>
<td>Residual</td>
<td>138</td>
<td>0.223333</td>
<td>0.001618</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>140</td>
<td>0.389632</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Coefficients**

<table>
<thead>
<tr>
<th></th>
<th>t Stat</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>12.29016</td>
<td>0</td>
</tr>
<tr>
<td>i - i$</td>
<td>-8.14313</td>
<td>0</td>
</tr>
<tr>
<td>VIX</td>
<td>-6.67485</td>
<td>0</td>
</tr>
</tbody>
</table>
Interest differentials were not large enough to compensate for the currency risk faced by yen carry traders who incurred loss, but the reverse yen carry traders enjoyed positive returns of 0.8 percent, 3.71 percent, and 9.93 percent returns on 1-month, 3-month, and 1-year short-term reverse yen carry trades.

Table 5 shows the results of regression analysis. This study’s regression model specification is found to be robust with highly significant F-statistic values of 13.43849 and 51.3787 with close-to zero p-values for 3-month and 1-year carry trades. R Square values are 0.02755, 0.136517, and 0.426809 for 1-month, 3-month, and 1-year reverse carry trade, respectively. Explanatory power of our two variables (interest differential and market volatility) increases as carry trade time horizon gets longer. Interest differential is negatively related to the return on the reverse yen carry trade for 1-month, 3-month, and 1-year, and the relationship is statistically significant with p-values of zero for 3-month and 1-year horizons.

Interest differentials, \( i_d - i_s \), are found to be negatively related to the reverse carry trade returns for all three time horizons. Japanese interest rates have been consistently lower than the U.S. rates during the sample period, that is \( i_d - i_s \) during the sample period is negative. For the reverse yen carry trade to be profitable nullifying the UIP, \( \beta_1 \), the coefficient for \( i_d - i_s \) is supposed to be negative, and it is negative for all three time horizons, thus positive returns on the reverse yen carry trade. These relationships are statistically significant with p-values of 0.024284 and zeroes for 1-month, 3-month and 1-year carry trade, respectively. Also, the statistical significance of the relationship between the interest differential and the reverse carry trade return increases as investment time horizon gets longer.

VIX appears to be negatively related to the reverse carry trade returns for 3-month and 1-year horizons, and these relationships are statistically significant with p-values of zero for 3-month and 1-year. The relationship between 1-month return and VIX is statistically insignificant. These results are consistent with my conjecture that: as market volatility rises, it becomes more perilous to engage in short-term speculative reverse carry trades that result in decline in return. An increase in the VIX by 1 would result in reduction of 0.151 percent and 0.181 percent on 3-month and 1-year reverse yen carry trade, respectively. Also, the statistical significance of the adverse effect of the VIX on the reverse yen carry trade return strengthens as the investment time horizon gets longer.
CONCLUSIONS

I investigated the relationship between the return on the reverse yen carry trades and two independent variables: Japanese risk aversion and LIBOR interest differentials, \( i_V - i_S \) for 1-month, 3-month, and 1-year investment time horizons, and market volatility as measured by VIX. Reverse yen carry trade returns were calculated with logarithmic transformation of realized spot exchange rates of yen: at time of yen carry trade, 1-month, 3-month, and 1-year after carry trade, during the sample period from October 2007 to August 2011.

I found that the reverse yen carry trades were profitable with statistically significant returns for 3-month and 1-year investment horizons as the interest rate differentials between Japan and the U.S. narrowed causing the commonly practiced yen carry trade to unwind leading to reverse yen carry trade under more volatile market condition. Volatility index (VIX) is found to be negatively related to the yen reverse carry trade returns. As market becomes more volatile, the return on the reverse yen carry trade declines.

LIMITATIONS OF THE STUDY

This study covers only the yen carry trade for the dollar for a limited period from October 2007 and August 2011. 46 weeks of “before” the 2008 global financial crisis compared to 146 weeks of “after” the crisis may be not long enough to infer any significant statistical relationship between the weekly returns on the yen carry trade and the two explanatory variables.

RECOMMENDATION FOR FUTURE RESEARCH

An explanatory variable of this study, “spread” as a proxy for the Japanese risk aversion exhibits a significant relationship with the carry trade return. This significant relationship sheds light on an increasing role of higher currency volatility in carry trade. It would be interesting to further investigate the relationship for the yen carry trades targeted for other currencies of higher interest rates than the yen by partitioning the impact of interest rate differential and increased currency volatility.
REFERENCES


THE SILVER TSUNAMI: EVALUATING THE IMPACT OF POPULATION AGING IN THE U.S.

Lauren Henderson  
Bala Maniam  
Hadley Leavell  
Sam Houston State University

ABSTRACT

The population of the United States is rapidly aging due to increases in longevity and decreases in birth rate, leading to what some describe as a “Silver Tsunami.” This demographic shift has the potential to create far-reaching consequences for the United States. As the overall age of the population increases, the number of individuals participating in the workforce declines, leading to potential economic decline if not countered with productivity increases. Further, increased demand for healthcare services could cause a dramatic increase in healthcare spending inflating an already massive government budget for Medicare. The demography changes also threaten the solvency of the Social Security program, causing policy makers to consider potential reform. Further policy reforms are considered to create a more age-friendly society such as, transportation, housing, and technology.

KEYWORDS: aging, population, demography, healthcare, Social Security

INTRODUCTION

The global population is aging precipitously, a trend mirrored in the United States. The cohort of US citizens over the age of 65 is predicted to more than double between 2015 and 2060, from 46 million to 98 million, creating a “Silver Tsunami” that predicated a number of challenges to the economy, healthcare, politics, and society (Mather, Jacobsen, & Pollard, 2015). A marked increase in longevity combined with a decline in birth rates has aged the global demography. While this demographic shift is more remarkable in other developed countries, it will still lead to pronounced changes in the demographic makeup of the United States. This older population is set to reduce the overall workforce and its productivity, increasing pressure on programs such as Medicare and Social Security, and in turn, generating added strain on the working age sector of the population. To support both the elderly cohort and the rest of society, the government will need to make dramatic changes to policies and institutions. In
addition, corporations and individuals will need to modify their behavior and decision-making.

The objective of this paper is to provide an analysis of the impact of this demographic shift within the United States, and to outline recommendations to counteract the negative consequences of aging, while enjoying the benefits of the collective longevity and experience of the elderly population. First, this paper will investigate the factors contributing to the aging phenomenon in the United States, with the global perspective presented as a point of comparison. The impact to the economy, the healthcare industry, public policy, and society will then be reviewed. Evaluating these subjects in conjunction helps to illustrate the interrelated nature of the demographic trends, and aids in providing comprehensive solutions. Finally, areas will be proposed where the research can be expanded to provide a more complete view of the ramifications of the aging society and the actions needed to prevent further negative consequences.

LITERATURE REVIEW

There are significant amounts of research and publications that describe the impact of population aging and potential solutions to counteract these challenges. Regarding the economic impact, the generally held view is that the labor force shrinks in an aging economy, which leads to productivity declines (Börsch-Supan, 2010). Eberstadt (2010) argues that the impact of a labor force reduction can be ameliorated through focus on innovation, investment in education, creating a positive immigration climate, and improving healthcare to create a healthier workforce. Some have argued that the potential economic shrinkage due to aging could cause a major solvency issue for the Social Security Administration. However, others argue that the solvency issue is overstated and that it can easily be corrected by investing in real productivity gains in the economy and providing incentives to join and stay in the workforce (Wray, 2006; Goss, 2010).

A major contribution to the economic outlook is related to healthcare spending and the increased demand for healthcare services. Hashimoto and Tabata (2010) demonstrate how the increased demand induces a shift to employment in the healthcare industry, and has the potential to impact the economy due to reduced productivity gains in the services sector. However, Aisa and Pueyo (2013) counter that this effect is ambiguous at best given that older populations increase personal savings to accommodate their increased retirement needs. This lends to an overall accumulation of capital, which tends to encourage labor productivity outside of healthcare, thereby improving economic growth.

In addition to employment shifts towards healthcare, the increased demand for healthcare suggests increased spending on healthcare, which constituted 17.6% of GDP in 2010 (Chinta, Burns, Manolis, & Nighswander,
Butler (2010) argues that health spending can be curbed through public health policies to reduce obesity, increase exercise, and provide universal health insurance. Reuben (2010) considers the ethical dilemma facing healthcare providers as they seek to ensure distributive justice for both the elderly and young who will compete for scarce healthcare resources.

Other research has centered around changes needed within the community to support an aging population with different needs from the current demography. Researchers have found a lack of accessible housing (S. Smith, Rayer, E. Smith, Wang, & Zeng, 2012). The World Health Organization (2007) proposes eight parameters leading to age-friendly communities. Hrostowski (2010) focuses on the benefits of “Aging in Place,” for both the senior citizen and the community where they reside. Rosenbloom (2010) critiques popular transportation solutions offered, as she finds that most elderly prefer to drive or walk rather than use public transportation.

To support the changing demographic, many researchers have proposed policy changes needed to enable current institutions to persist and thrive. Many argue that a positive immigration climate and incentives can offset population aging, and avoid economic decline (Calahorrano & Lorz, 2011; Kotkin & Ozuna, 2012). Others have argued for reform to Social Security and Medicare (Munnell, 2010; Wray, 2006). The aging voting populace promises to keep these issues top-of-mind for politicians, but studies have shown that the depiction of elderly in popular culture and in the media, can have an impact on the nature and partiality of legislation passed (Keyes & Dicke, 2016).

While the research on the impacts of the aging population and potential mitigating solutions is vast and well-documented, there is a need to provide more comprehensive solutions that consider the interconnectedness of these different policies. As healthcare spending can impact the economy, solutions intended to correct for that spending can have far-reaching consequences if they are not considered in tandem. This paper will consider the interplay of the economy, healthcare, public policy, and society in the United States.

**HOW IS THE DEMOGRAPHIC MAKEUP OF THE UNITED STATES CHANGING?**

Developed countries are experiencing population aging to an extent that has never been seen before. “Population aging” refers to the percentage of individuals over the age of 65, in comparison to those under 65. There are three perspectives on aging – individual, population, and global (Lee, 2014). From an individual perspective, life expectancy in the United States has increased from 68 years in 1950 to 79 years in 2013. Individuals are also more likely to spend a greater part of their lives disability-free due to advances in medicine which have improved the prevention, early identification, and treatment of disease (Mather et
Meanwhile, the overall fertility rate in the United States has declined among native citizens from 4.9 births per woman in the early 1960s to an estimated 2.5 births in 2010 (Eberstadt, 2010). The immigration climate in the United States has helped to off-set the lower fertility rates as compared to other developed nations (Kotkin & Ozuna, 2012). However, the collective effect of increased longevity and decreased fertility results in population aging in aggregate.

Globally, population aging is occurring in an even more pronounced way than in the United States. Nearly half of the global population lives in countries where the fertility rate has fallen below the replacement level, leading to rapid population aging and even population decline. (Eberstadt, 2010; Wray, 2006). By the year 2050, it is estimated there will be more adults over the age of 60 than children under 14 years old for the first time in history (World Health Organization, 2007). This shift in the individual, population, and global demographic makeup is likely to have significant impacts on economies and policies. While the specific implications to the programs and citizens of the United States are less clear, observing other developed countries provides a chance to forecast the impacts within the United States.

The Baby Boomers Generation symbolizes the pronounced shift in the demography of the United States. The Baby Boomers are the 76 million people born from 1946 till 1964 (Mather et al., 2015). This generation is the result of peak fertility rates following the conclusion of World War II, resulting in a larger cohort than generations on either side. This group has come to represent the trend of the aging population, and has been much maligned in political discourse for threatening to break the bank on programs such as Medicare and Social Security.

While the Baby Boomers will not single-handedly cause the demographic shifts described globally, the retirement of the group will amplify its consequences and create added strain on already fragile systems. The oldest of the Baby Boomer generation will reach 80 years old in 2026, creating an urgent platform for change in public policy and healthcare. Changes made to policies and programs to stabilize them during this period will ensure the permanence of the programs for future generations who will demonstrate the same demographic trend, albeit less pronounced.

THE ECONOMIC IMPACT OF AGING POPULATION

One of the most obvious and consequential outcomes of the aging of the population is the economy of both the United States and all global nations. The overall birth rate and immigration rates in the US have declined since 1950 which in turn causes a decline in the labor supply which could lower the overall productivity of the US Economy. This trend is predicted in most developed nations.
The United States has experienced higher fertility rates than other countries in recent history, which means this labor force reduction may not be as extreme (Börsch-Supan, 2010). However, the labor force will still be older on average. There is an assumption that older employees are less productive, due to resistance to technology and innovation, and diminishing health (Eberstadt, 2010). If the US is to prevent economic decline and maintain its status as a global power, it will need to identify ways to both bolster the working age population and increase the productivity of older workers.

One way to buttress the diminishing labor supply is by encouraging immigration and ensuring that immigrants to the US are integrated into society and successful in attaining education and employment opportunities (Kotkin & Ozuna, 2012). Another strategy is to improve the overall productivity of the workforce with a focus on technological innovations such as automation and artificial intelligence to enhance the efficiency of each worker. Companies may also consider adding flexible work schedules and alternative work arrangements to accommodate older employees desire to balance work with leisure. This flexibility would encourage older adults to stay in the workforce longer. It is also critical to improve healthcare so that more of the labor force is disability-free and capable of working.

On the flipside of the coin is an increase in consumption of goods and services that is found in both the very young and the elderly. The Old Age Dependency Ratio (OADR) is one way to measure the economic impact of the aging population as it measures the relative impact of dependent elderly to younger adults. The OADR is projected to rise from 0.22 in 2010 to 0.39 in 2050 (Lee, 2014). This marked increase in the OADR demonstrates the increased demand that will be placed on young adults to support the greater needs of the elderly dependents. The demand model for an individual predicts increased demand at two points in life – as a dependent youth for support and education, and as a senior over the age of 65 for healthcare and retirement.

Some argue that increased demand for healthcare services will cause an overall shift in workers from other industries to the healthcare industry overall (Aisa & Pueyo, 2013). Since the working-age population is set to decline overall, this shift of workers to the healthcare industry will cause the ratio of workers in the healthcare industry to increase relative to other industries. The services sector of the economy (including healthcare) has not exhibited productivity gains to the same extent as goods sector of the economy, and productivity increases within industry is believed to be the primary source of economic growth (Hashimoto & Tabata, 2010). Therefore, it can be argued that the increased ratio of healthcare to non-healthcare industry in the economy will decrease the overall economic growth in the United States. However, this factor may be minimized or offset by the increase in capital accumulation that accompanies increased longevity (Aisa & Pueyo, 2013).
Another significant impact to the economy is an overall reduction in savings that accompanies the aging population. Throughout their working life, individuals accumulate savings to prepare for retirement. After retirement, seniors begin to disinvest their retirement funds, resulting in an overall reduction in savings (Börsch-Supan, 2010). Individuals are net savers and the government and corporations are net users. If the overall amount of savings declines, the amount of money available to the government and to firms shrinks, which will drive up interest rates. Since the interest rate is a key driver of the weighted average cost of capital, it will inevitably impact the financial decision-making of businesses. While logically this argument holds true, observations of Japan’s elderly population have shown the assumption that saving behavior decreases in old age to be invalid. Further research is needed on saving and consumption patterns the U.S. elderly to fully understand the impact of savings to the American economy.

Americans are retiring at an average rate of 10,000 per day - 80 million over the next 20 years - and spending more years of life in retirement (Social Security Administration, 2011). This leads to another economic challenge - how to support retirement income for older individuals. Retirement income is typically funded through three vehicles – individual savings, family bequests, and government programs (i.e. Social Security) (Munnell, 2010). Recent observation shows that contribution to income in all three areas is declining.

First, there has been a shift from guaranteed monthly income via pensions to individually funded 401Ks. Limited financial education has left individuals lacking understanding of how to properly prepare for retirement. This means people do not effectively use this vehicle – investing less than the maximum, holding minimally diversified portfolios, and choosing riskier investments (Munnell, 2010). In addition, when individuals start withdrawing from their 401K, they devalue annuitized plans, leading to faster depletion of funds. The second, albeit smaller contributor to retirement income is family inheritance. This is also declining, as increases in longevity lead to depletion in family coffers (Gokhale & Kotlikoff, 2001).

The final means of retirement income in the U.S. is arguably the most important – Social Security. Sixty percent of retirees rely solely on Social Security for retirement income. With the aging population, the number of workers contributing via payroll tax is declining while the number of individuals withdrawing is increasing. The Social Security Administration predicts that program costs will rise to the point that, by 2035 taxes will cover only 76% of Social Security benefits (Goss, 2010). This has led many analysts to contend that Social Security faces a major solvency problem, meaning the trust will be unable to pay full benefits on a timely basis. Some argue that this position is unfounded, and that by focusing on real economic growth the Social Security Administration will avoid any sustainability concerns (Wray, 2006). The US government will need to either increase contributions by incentivizing work and immigration, or decrease
withdrawals. There is also an optics issue at hand: the public views Social Security as a pay-as-you-go benefit, when in reality it is a tax and insurance benefit where they will not receive dollar for dollar what they invested (Wray, 2006).

In summary, all individuals will need to work longer, save more, and efficiently use savings while the government will need to increase the retirement age, reform entitlement programs, and improve financial education in order to sustain retirement income for older generations (Brooks, 2010). Healthcare comprises nearly 18% of GDP spending and 40% of government spending, representing a very significant sector of the US economy (Chinta, et. al., 2013).

THE AGING POPULATION IMPACT ON HEALTHCARE

As the population ages, there are growing rates of physical and cognitive decline leading to an increased demand on all types of healthcare services. While life expectancy has increased significantly due to advancements in public health as well as in the treatment of disease, experts predict that life expectancy in the US may decline due to rising obesity, chronic diseases, and lifestyle choices such as alcohol and tobacco use. Americans may live longer but sicker lives, increasing the burden on the healthcare industry. To combat rising healthcare costs, programs focusing on promoting healthy lifestyles through proper nutrition, adequate exercise, and reducing alcohol and tobacco use should be expanded. The government should also consider consumption taxes on unhealthy choices (e.g. soda tax) and improving health insurance to promote preventative care (Butler, 2010)

Another contributor to the current healthcare crisis is the payment model. Healthcare providers are incentivized in the current model to perform aggressive surgical procedures and recommend hospitalization, penalizing providers who use lower acuity treatment alternatives with low reimbursement and relative value units (RVUs). This model resulted in skyrocketing costs for healthcare in the United States; in 2005 the US spent $6,041 per capita on healthcare, representing twice the per capita spending of other industrialized countries with similar quality (Anderson & Frogner, 2008). With the Affordable Care Act, new payment models are being tested that incentivize quality outcomes and cause healthcare organizations to carry the financial risk of their treatment outcomes (Jones, 2016). This trend is known as “Value Based Care” or “Bundled Payments,” and it is driving firms toward lower cost treatments. These payment models may help address the shortage of primary care providers, and specifically geriatricians, by making the compensation more competitive with specialists.

The rising cost of healthcare has been a primary concern for individuals and politicians alike, as the government pays for 40-50% of healthcare costs through the public insurance programs Medicare and Medicaid. The dramatic increase in chronically-ill Medicare patients will precipitate a rapid rise in
healthcare spending by the government, stressing an already stretched budget. The best approach to reducing healthcare spending is to focus on promoting healthy lifestyles—especially early in life. Early education breeds behavioral changes that produce increased dividends later in life (Kramer, 2010). However, research is needed as to what drives individual decision making and compliance with healthy lifestyles. Researchers to date have been unable to determine the factors that lead individuals who understand the risk of behaviors of sedentary lifestyles, alcohol consumption, and smoking to continue with those destructive behaviors.

The increased demand on the healthcare system, and limited supply of financing and healthcare providers raises an ethical concern of whether to spend $16,000 on a week-long ICU stay for a patient at end-of-life when that could pay for a year of health benefits for a family of four. Mechanisms and protocols are needed that create an incentive for the right decision, rather than allowing families to wait on miracles for patients at end-of-life (Reuben, 2010). Without these tools on hand for healthcare providers and organizations, the increased healthcare demand promises to exhaust both physical and financial resources to provide high quality care.

Finally, additional funding is needed to help address chronic diseases that manifest at greater rates in the elderly such as cancer, heart failure, and diabetes. Chronic, incurable disease is extremely costly, not only financially but in reduced ability of those individuals to contribute to the workforce and society. Research is also needed to investigate prevention of “frailty,” a loosely defined set of conditions that foretells decline and mortality. Early evidence shows that frailty can be reduced or prevented through active engagement of seniors in their community, accessible housing, and technology (Fried, 2010).

SOCIAL IMPACT OF AGING POPULATION

The way society portrays age in daily conversation, and in the media, impacts the real actions by society to support the elderly. If the population consistently hears negative, dependent stereotypes, they are unlikely to value the contributions of the elderly and will not be willing to contribute towards those causes (Keyes & Dicke, 2016). Society needs to emphasize the fact that improving lifestyles of senior citizens really benefits the whole family, as elderly relatives can contribute to child care, or generally contribute the wisdom and experience of their years (Fried, 2010). It is also important to note that supporting the elderly should not start at age 65, as healthy lifestyle promotion should begin early.

The World Health Organization (2007) outlined eight dimensions that help to define a community best suited for geriatric adults—outdoor spaces and buildings, transportation, housing, social participation, respect and social inclusion, civic participation and employment, communication and information, and community support and health services. One of the basic, but specialized needs
of seniors is housing. Most seniors require some sort of assistance with Activities of Daily Living (ADLs) as they get older. Rather than receive this assistance in a nursing facility, there has been a recent trend towards “Aging in Place,” which has boosted the demand for community based care such as home health agencies, adult daycare, and community care facilities (Sullivan, 2016). This has been supported by provisions of the Affordable Care Act, which provided financial incentives to states that support these programs. Job growth in these areas promises to shift some of the available resources to these opportunities, but must be considered carefully as there is a potential to reduce economic growth due to the reduced productivity growth in the services sector (Aisa & Pueyo, 2013). There are great benefits to seniors able to stay in their home or in the neighborhood, it creates a sense of community that increases social engagement and has been shown to prevent cognitive decline. It also promotes inter-generational connection and prevents the sort of ageism described above by allowing seniors to volunteer and contribute to their society.

Homes with accessible features such as wide doorways, zero stair entrances, and first floor bathrooms and bedrooms are needed. A remarkably small percentage, 10%, of U.S. residences have all these features, and many of them can be credited to The Fair Housing Act Amendment of 1988 requiring apartment complexes to include these amenities (Smith et al., 2012). Creating communities that better support older adults provides an avenue for a greater part of the elderly population to stay in their neighborhoods. Having a neighborhood with a mixed age group present can improve both quality of life of the individual residing in the neighborhood and the quality of life for the entire community (Hrostowski, 2010).

Another basic need of the elderly is access to transportation. Nearly 20 percent of the elderly population lives in rural areas. This geographic isolation is attributed to widening health care disparities noted in the population (Martin & Finch, 2016). Transportation is critical for these adults to access life-sustaining care, but also to allow them to stay socially connected. Most elderly adults use automobiles as their primary means of transportation. Studies show that they have accident rates higher than average adults, but less than teenagers. Infrastructure needs to be improved to support safe driving - for example with lane design and signage (Rosenbloom, 2010). While some studies have suggested a buildup of public transportation systems, others argue that a very small percentage of seniors choose to use these forms of transportation (Hrostowski, 2010; Rosenbloom, 2010). In practice, when seniors cannot drive, they are more likely to walk rather than take public transport. Walking not only provides a form of transportation, but also provides health benefits as a form of exercise. Local governments need to improve the safety of walking by providing islands in crosswalks to extend crossing time for mobility-impaired seniors, enforce limits on cyclist and skateboard traffic on pedestrian walkways, improve lighting and policing of walkways, and provide adult-oriented public spaces to promote walking.
This paper already explored the impact of increasing healthcare demand on the demand for healthcare employees. In addition to the professional caregivers in the healthcare industry, many individuals will serve in informal roles as caregivers. Typically, this role is provided by the adult children or other family members of the senior. In recent years, the number of caregivers has been declining relative to the older population (Mather et al., 2015). The government and businesses need to provide better support to informal caregivers, as they are instrumental in keeping seniors in their homes and out of institutions. Potential technological advances in remote monitoring and telemedicine may make it easier for caregivers to continue to work while they provide care; this would provide benefits to the seniors, but also boost the economy by keeping workers in the workplace (Coughlin, 2010).

Another essential need of older adults is socialization. While this seems a lower-order need, there are studies that show that seniors actively engaged in their social circles stave off cognitive decline and show greater emotional wellbeing. As mentioned previously, transportation can be a concern for seniors, so technology may be a suitable alternative to support integration of the elderly person with their community and family virtually (Coughlin, 2010). Social networking designed specifically for seniors, and computers designed specifically for seniors with physical limitations is likely to grow in popularity.

The integrated nature of the challenges of population aging is evident, as the US will need to address nearly all aspects of the economy, healthcare, and social programs to prepare for the demographic shift. Some of these changes will take place within individual organizations, but much of the change will come through government intervention.

**IMPACT OF AGING ON PUBLIC POLICY & OTHER ISSUES**

One of the biggest public policy struggles is the “Structural Lag” in US institutions such as Medicare and Social Security. The government must find a way to improve responsible policy-making based on demography changes (Moen, 2010). Policy decisions are increasingly interconnected as energy policies impact environmental and therefore healthcare spending (Swartz, 2010). It is critical for policy makers to consider comprehensive reforms to address the global changes in the economy and environment. Research has shown that health care costs are not consistent between regions or even within age demographics (Chinta et al., 2013). However, further analysis is needed to understand how to shift the curve towards higher quality and lower cost care.

Additionally, this analysis has explored ways to prevent a solvency issue with Social Security. Public policy should focus on incentives to work, such as decreasing the payroll tax in favor of supporting Social Security via progressive property tax (Wray, 2006). The government should also set about improving the
financial literacy in the US, helping individuals to better prepare for their future financial needs (Brooks, 2010). Likewise, investment in education for the younger generations promises to boost their labor productivity and spur wage growth that could help to improve incomes for the Social Security trust which could alleviate the strain of increased withdrawals (Stauvermann & Kumar, 2016).

Older adults are more likely to stay engaged in politics and in the voting booths. With the percent of the population shifting towards the elderly, this means that issues impacting the elderly will become more pertinent to political discourse. Some speculate that there will be a shift towards conservatism, as elderly hold more traditional views on immigration, family, and economic issues. Immigration affects generations differently, as older generations show increased incomes from boosts to overall economy, where younger generations may suffer from increased job competition (Calahorrano & Lorz, 2011). Politicians will need to balance the current needs with the needs of future generations. They should also emphasize the benefits that supporting elderly provide to the greater community. More research is needed to delineate the benefits of these programs to the larger society.

Another policy consideration to be addressed is immigration. Strong immigration rates to the US can boost economic growth. In addition to sustaining population growth by boosting the birth rate and augmenting younger age brackets, immigrants also show more entrepreneurial initiative than native populations. Immigrants have played a significant role in boosting small business within local communities in the US. The current political narrative is counter-immigration and has the potential to handicap the future economy (Kotkin & Ozuna, 2012). Education to the general population espousing the benefits of immigration and explaining the need for importing skilled labor is essential. Updating immigration policy to promote skilled labor migration to the United States in order to specifically target the technology and healthcare sectors would be beneficial (Calahorrano & Lorz, 2011). Policy makers should also expand the education and governmental systems to support successful integration of immigrants.

**IMPLICATIONS AND FUTURE DIRECTIONS**

From an economic viewpoint, labor supply, productivity of the labor force, changes in demand, and saving patterns are of concern. With many developed countries already experiencing the rapid aging phenomenon, economists should study the impact and extrapolate to the United States. For example, by observing Japan who has experienced an incredible workforce decline, economists can understand what policies have helped to mitigate economic deterioration. They might also seek to understand from the cohort nearest retirement what programs might encourage citizens to stay in the workforce to provide firms with greater direction in creating flexible work programs.
It is also critical that the US population be educated on managing personal finances to ensure that they are prepared to support themselves in retirement. Scholars must devise ways to communicate this information to the layperson, using innovative delivery mechanisms such as social media. Additionally, legislators must create an open dialogue about amending the social welfare programs in the US. Rather than seeking to make sweeping changes, policy makers should work to create a political environment that supports iterative legislation that can be implemented quickly and adjusted based on the response.

Researchers have detailed the rising healthcare costs and the inequitable costs between healthcare providers. Further analysis is needed to identify the cost drivers and determine how to provide the same quality care at a lower cost. Simultaneously, healthcare organizations should open dialogue on distributive justice in healthcare. While some persons may view it as a taboo topic, researchers predict that the US will face healthcare shortages and it is critical to approach those shortages with an impartial strategy, rather than reaction from irrational emotions.

Another particular area of focus must be the promotion of overall health and wellness, both within the elderly and within the general population. Nearly 80% of disease outcome is driven based on lifestyle choices such as nutrition and exercise, but very little is known about what specifically determines individual decision making. Further research is needed to determine how healthcare providers can influence behavior change and make lasting differences (Kramer, 2010). The healthcare industry must also address the shortage of primary care providers by studying what programs would incentivize new doctors to the fields of family medicine and especially geriatrics.

There are early adopters in the US practicing some of the community-building techniques proposed by researchers to improve longevity and productivity of the elderly population. In-depth analyses should be completed on these programs to understand the true impact, providing data to fuel future innovation. Additional evidence needs to be collected and analyzed to describe the positive benefits of keeping elderly within their communities, as this information may persuade the general population to support community-based housing for seniors.

Finally, further research should be considered to understand the complete view of policies and actions to prevent negative impacts because of and for the aging population. There is a need for comprehensive and longitudinal analyses of programs on the full population. For example, if the US focuses on improving early childhood education on nutrition, does that reflect in increased longevity and prolonged productivity in older adults (Butler, 2010)? Emphasis on the interplay of policies in different spheres will deter isolated departments from issuing legislation that is well-meaning but has unseen consequences in other areas.
Although much has been written and conjectured about the impact of an aging society, there is still myriad topics that require further review. To support research in these areas and encourage innovative thinkers to address the challenges of population aging, the United States will need to allocate funding towards these programs. Researchers will need to both study the impact of aging. Advocates will then be needed on behalf of the aging population and the nation as a whole to acquire the resources needed to explore comprehensive and longitudinal studies.

CONCLUSION

The impact of the aging population will be significant to the United States if society does not adjust behaviors and institutions to support the new demography. Without intervention, the labor force will shrink, resulting in minimal or declining economic growth. This will create increased strain on the existing workforce to support the older retired population. This trend can be diminished by improving the incentives to work, both for elderly and for working age immigrants. Companies should also aim to increase actual production of goods and services through efficiency gains, and a focus on technology and innovation. Politicians also need to address the structural issues in the social programs of Medicare and Social Security. While intended to serve as a safety net, they increasingly represent the sole source of income and insurance for the elderly population. These programs need reform to safeguard solvency and to reduce overall costs to permit sustainability. Local institutions and individuals in society will also need to prepare for the aging to stay present within the community. The transportation, housing, and social needs of seniors need to be considered when planning community projects ranging from infrastructure to environment.

The “Silver Tsunami,” has the potential to generate serious issues in many areas of US society – economy, healthcare, social programs, and public policy. However, these concerns can be lessened or eliminated through comprehensive research and reform. Everyone, from politicians to physicians, will need to prepare for and address issues within their spheres of influence. An aging population does not imply disaster, rather it presents an opportunity for the citizens of the United States to band together and provide comprehensive reform to benefit all.

REFERENCES


Henderson, Maniam and Leavell


Henderson, Maniam and Leavell


INCENTIVIZING MANAGEMENT DISCRETIONARY PHILANTHROPY: SOCIAL PROFIT CREDITS

Marcel C. Minutolo
Chloe Mills
John Stakeley
Robert Morris University

ABSTRACT
Discretionary giving remains an important tool for corporations to demonstrate social responsibility and reflect corporate values. Funding mechanisms within nonprofit sector are under stress from recent broad economic upheavals combined with a lack of resource diversification. In order to offset the current inefficiencies, the authors present a conceptual model of Social Profit Credits (S\pi C) as a mechanism to encourage and enable competition and innovation in the nonprofit sector that more closely reflect traditional capital markets. The S\pi C builds on existing models such as social impact bonds (SIB), carbon credits and renewable energy credits, and extends those to the nonprofit domain. In combination with appropriate sector-dependent evaluative techniques, the S\pi C can encourage the activities necessary to support nonprofit missions, reduce dependency on restricted grants, diversify funding resources, and increase overall social benefit while incenting firm due diligence. The value of the current work is to provide a mechanism whereby corporate discretionary giving is applied to increase overall nonprofit performance resulting in increased social benefit.

Key Words: Social profit, strategy, nonprofits, philanthropy, competition

INTRODUCTION
Nonprofit organizations (hereafter nonprofits) are voluntary service providers who serve a philanthropic social need that traditional capital markets have failed to meet. Like their traditional capital market counterparts, every nonprofit aligns its activities with a mission geared to a specific task (e.g. veteran needs, literacy, nutrition, etc.). Where the commercial enterprise is able to generate revenue through the sale of its services or products, nonprofits raise capital through solicitation and collection from individuals, businesses, and the Government agencies. As reported by Ozdemir, Altinkemer, De, and Ozcelik (2010), the nonprofit sector, with nearly a trillion dollars in assets, accounts for five percent of the United States gross domestic product. Additionally, the average American donates more than two percent of his or her annual income to nonprofit causes (Ozdemir et al. 2010).
Despite the overall importance to the economy, most individuals and organizations focus their discretionary giving not on the performance of the organization but rather on the appeal or the emotional connection to the cause. Further, grants provided by the Government are typically awarded based on a narrative rather than by demonstrable results; end of grant performance evaluated by the amount of the grant that went to direct service regardless of performance. Nonprofits generally have a great deal of difficulty providing documentation for their performance because it is not always clear how to measure the benefits that these entities provide to society. That is, while a nonprofit organization may be able to point to some quantitative information, such as the number of families serviced, it is not always certain that the means by which the service provided is as efficient or even as fully beneficial as the organization might hope. Additionally, given the difficulty in evaluating the performance, many nonprofits have resorted to merely evaluating the relative amount of donations that go to service as opposed to overhead. The challenge with this model is that it also limits the ability of the organization to perform many of the functions that could increase efficiency.

When considering the value chain of an organization, all of those items that occur in the secondary activities (R&D, HR, Facilities, Procurement, IT) are reduced to a small percentage of total revenue, generally 10% under the commonly used 90/10 rule-of-thumb, by which nonprofits are said to be performing well if 90% of revenue goes to the provision of program services. This percentage can be difficult to maintain. Additionally, it can limit the organization’s ability to invest in capacity building activities that may not have a short-term demonstrable impact but are put in place in order to defend a position, anticipate future demand, or enter into new market spaces. However, given the constraints placed on the organization to ensure that most of the funds go directly to service, it is difficult too for the organization to innovate and engage in these growth activities. Therefore, nonprofits need a different source of funding that provides them with the freedom to manage their organization unencumbered by spending restrictions to better serve their populations.

With more than 1.7 million nonprofit organizations in the United States listed in GuideStar (www.guidestar.org), it is extremely difficult to gauge which of the nonprofits a firm should allocated its discretionary giving toward. In fact, there is little incentive beyond the tax deduction and goodwill for firms to spend too much managerial effort toward the decision making process. Hence, to better align the interests of the nonprofits and philanthropic giving, a better mechanism beyond tax deductions needs to be created.

As stated, nonprofits exist to serve social needs where traditional capital market are unwilling to step in. While there may be no clear profit mechanism for these initiatives, society still values the services that these organizations provide and authorizes the allocation of resources to meet the needs that the nonprofit organization fills. However, grants and donations are increasingly scarce and,
when available, are often restricted. Therefore, the nonprofit sector needs to diversify funding sources in order to succeed and thrive.

In this paper, we suggest the creation of a funding mechanism we call the Social Profit Credit (SπC) that has the potential to better benefit the market, better serve the nonprofit sector and its population, and increase overall social profit. This credit is an extension and elaboration of the social impact bond (SIB), a funding mechanism enjoying some recent public attention (Azemati et al., 2013; and, Baliga, 2013). We demonstrate how the SπC can work to improve performance, stimulate innovation, and act as a mechanism for an efficient market for nonprofit investment.

SOCIAL PROFIT CREDIT

Social profit (Sπ) is a conceptual proposal for evaluating the “humanitarian benefit gained as the result of investing in the well-being of others” (Gilligan and Golden, 2009) of both nonprofit organizations and for-profit firms that support some kind of social mission apart from the bottom line (e.g. Ben & Jerry’s, Newman’s Own). In essence, Sπ is an attempt to quantify the good that is done by these types of organizations. To further define the concept of the SπC, let us look at the means by which different types of institutions are funded within society. There are, essentially, three different dimensions of the socio-economic fabric. These dimensions include: for-profit (FP) institutions, nonprofit (NP) institutions, and governmental organizations (GO), which utilize such funding mechanisms as the market economy, corporate philanthropy, social investment bonds (SIB) and individual charitable donations, grants, and tax revenue.

Each institutional region is dominated by different modes of transaction. The types of activity in overlapping regions are bilateral. That is, region one will include at least tax revenue for the government and various tax credits to encourage capital markets. Region two includes government grants to NPs, and the economic value of the NP provision of activities for which no capital market exists. Finally, in FP philanthropy resides in region three; for their part, NPs provide FPs the opportunity to gain economic offsets and goodwill that can be accounted for on the firm’s books. Individual philanthropy constitutes transactions which appear in each of these activity regions.

In current practice, the central area of overlap of all mechanisms of resource-provision and institutions is unfilled. We propose that the currently empty region of institutional and economic overlap be filled by the Social Profit Credit (SπC). The extension of the current work is reflected in the proposal of the SπC as a funding mechanism that frees nonprofits from restricted funds resulting in innovation, mergers and acquisitions, and competition that more closely reflects traditional capital markets.
FUNCTIONALITY

In order to fully encourage market dynamics that will improve the overall performance of the nonprofit sector, extending the conceptual models of SIBs and develop a financial instrument that is more similar to carbon markets. Like SIBs, in order to successfully negotiate, issue, and contract an SπC, different parties have to be involved and support the undertaking. The government, nonprofit, and investor all have to be aware and agree on the terms outlined in the SπC (Callanan et al., 2012). The creation of an exchange for the SπC to be traded needs to occur; we envision that this could be accommodated in existing marketplaces such as the NYSE or other large exchange. Third-party evaluators need to be established and vetted. Again, we believe that this can be accomplished through existing mechanisms. One likely source for third-party evaluators may come from auditing firms. Finally, there needs to be a source for the underwriting of the credit itself. The agencies that print the credit may take the form of existing organizations that already underwrite bonds and other issuances. Unbiased data analysts are also required to track the nonprofits progress during and at the close of the SπC (Callanan et al., 2012).

To illustrate how the SπC works, we suggest the following steps to establishing the actual SπC. A nonprofit agency engages with a third party organization (hereafter Agent) whose role it is to take the credit to the market for sale. The Agent has several requirements in this process. The first expectation is that at the beginning of the period under consideration, the Agent, working with the NP determines what the appropriate accounting metric ought to be; this is generally established by the sector within which the organization works. After the organization engages with the Agent, it is necessary to establish and calculate the social profit metric with respect to the relevant nonprofit sector. Social profit constitutes the abnormal social returns achieved beyond the sector average. In any sector, the denominator ought to be the same: the total cost of delivering the service. The total cost of delivering the service ought to be easy to determine by merely looking at an organization’s 990 report. The numerator, on the other hand, may be much more difficult to determine since there may not be an agreed upon metric (see equation 1). In the case of providing meals, it may be the number of meals provided or the number of individuals fed.

The SπC may be calculated as follows:

\[
\text{Equation 1: } \text{SπC} = \left[ \frac{1000 \times \text{Percentile}}{1 + r_f + r} \times (1 + r_c) \right] + \text{Fee}
\]

Where \( r_c \) is the risk-free rate; \( r_f \) is the write-off rate based on previous performance, \( r \) is the credit premium for expected performance during that period; and, \( \text{Fee} \) is the associated transaction costs. Percentile ranking is used as an overall measure of individual risk. The lower a particular organization ranks with respect
to the target metric relative to the others in the market space, the lower the believed probability of achieving the outcomes in the coming year. The one-year risk-free rate captures the minimum rate of return that an investor expects to realize from the instrument. The risk premium for SπC is the inherent potential for failure, this is the measure of risk used to discount the SπC for exposure. Finally, commission and fees are the amount charged by third parties for the transaction.

These four elements determine the pricing of any individual issuance that an investor would pay for a credit. The face value of each credit may be issued in $1,000 increments. However, in a competitive environment, based upon variable rates of return, receiving face value for the SπC is risky. Receipt of face value or greater implies that there is no risk and that the expected performance is high enough above average to offset commissions and fees. Each organization then will exercise managerial discretion to decide on what dimension they have the resources and capabilities to compete. Like traditional markets, some firms will attempt to control the variable in the numerator; some the denominator; and still others, will try to compete on both dimensions.

The competitive environment that ensues from this framework continually pushes performance up such that even the best firms need to achieve higher levels of performance in coming years since part of the calculation is determined by sector averages. A top tier SπC allows an investor to write off an amount very close to or perhaps even greater than the face value of the instrument. In this way, the SπC is an instrument that has the potential to return a traditional tax write off of the face value plus the credit for abnormal performance. In all cases, the firm is able to write off the initial investment less commissions; in this way, the SπC mirrors investment in stocks.

At the end of the year, the agent is responsible for certifying actual performance. The write-off value is determined for the credits of each institution using the industry approved metrics. It is important to note that because these credits have terms for only one year, it is imperative that all credits are sold at the beginning of the year and data is submitted immediately after the end of that year. While we can image that there is a secondary market that trades credits during the year, the resale price of the credit is dependent upon new information in the market and the revenue realized from the resale is not realized by the social profit agency unless it withheld some form the initial offering for future sale.

Some points to note in formula 1 - the Sπ, as mentioned, is the difference between the total cost of delivery and the reach. The organization then can choose to compete on one of several dimensions in the same what that the for-profit counterparts do. On the one hand, the NP can try to improve performance thereby reaching more or providing better service depending on how it is measured in the sector: a differentiator. On the other hand, the organization could attempt to compete on cost thereby delivery the service for less while maintaining the same
level of reach. The goal is to improve the margin; those organizations that have a greater margin are said to create $\pi$

The $\pi$C is issued with a par value of some amount. For sake of discussion, we will state that the par value of a fictitious $\pi$C is $1,000. The issuer of the certificate takes a commission; assumed 10-percent for discussion. Since there is no guarantee that the organization is going to achieve the level of performance stipulated in the credit, the purchaser buys the credit at a discount equivalent to the expected level of risk.

Our model, adapted from Husted and Allen (2009), illustrates the differences between traditional, FP, markets and the NP market. Plotted on a graph, the $y$-axis may be monetary units and the $x$ axis the reach, the measure of which will depend on the sector. For a rational firm in the FP market place, the organization will stop when the tangents of lines benefit and costs are parallel and intersect the benefit-cost curves; the point where marginal returns from investments begin to decrease. However, the traditional NP organization who is not burdened by meeting investor expectations in the same way will continue to invest till line three; the point just before one addition dollar of investment has a loss. The total amount of social gain is calculated by determining the total area between the cost and benefit curves.

There are firms in the SP market that create over-performance relative to the sector average. Firms that create social profit beyond the sector achieve abnormal social profit which is achieved when an organization is able to 1) push the benefit curve up; 2) push the cost curve down; or, 3) both. The result of moving the benefits and/or cost curve is to create additional area within the cost-benefit region. The “credit” then that the firm will receive for the investment instrument is the excess area within the $\pi$ region.

Conversely, not all firms perform the same. An organization that is 1) unable to produce the same level of benefits per dollar of investment relative to the sector average; 2) unable to keep costs as low as the sector average relative to the same benefit reach; or, 3) is unable to achieve sector averages in either costs or benefits.

A particular note is that the under-performing SP firm still creates SP, just at a level less than that of its sector mates. Equation one captures the continuum of differences in performance for all SP firms in a particular NP realm. Theoretically, one could plot the cost-benefit curves of each SP firm, but equation one allows the market to capture this relative performance for any one firm. The risk that the investor of the instrument is exposed to is the result of purchasing an instrument of a SP organization that is unable to achieve expected levels of performance. Therefore, the investor is only able to write-off the loss and does not realize the credit.
In the proceeding section, we illustrated how the SπC may be calculated and the steps where it may be issued and brought to market. Like a bond, the initial issuance has to be large enough to justify the investment in the creation and sale which implies fairly large orders of magnitude; this may not be viable to small social profit organizations. Further, we depicted hypothetical levels of performance relative to sector averages and how the social profit may be visualized as well us under performances. In the following sections, we discuss the benefits that all stakeholders may realize by moving to the proposed investment instrument.

**BENEFITS FOR NONPROFITS**

Among the outcomes of the proposed investment instrument, one we hope to see to fruition, is a relabeling of the nonprofit sector to the social profit sector. This concept helps to clarify through terminology the very reason for a specific NP to exist, namely the provision of an unqualified social good. Additionally, in pulling all entities that provide social profit together, the nomenclature highlights the community of need and shows the public that such organizations are working for them, rather than for some undefined other. The notion of “nonprofit” only underscores a specific fact of an organization, namely that it is not intended to make money for its shareholders. Whereas the terminology of “social profit” indicates that the deliverable of these organizations is social, common, and shared. Nonprofits, constrained at times by somewhat old fashioned thinking about what they are doing and whom they serve, can use the concepts of social profit to enhance their own stories, and their missions. In a funding environment which is increasingly competitive, a well-articulated social profit message will help social profit organizations gain the investment they seek.

Not all nonprofits will choose to participate in the social profit credit. However, we envision a cascading effect (Hess and Dunfee, 2007) that will lead to greater levels of participation in the SπC market over time since those organizations not participating will become noncompetitive. The early players in the market are likely to be the ones that are already performing at higher than average levels with respect to the service that they provide. The higher achieving firms will want to leverage their performance in order to get access to greater levels of unrestricted capital to continue to fund their activities. They will quickly see their portfolio of revenue diversify and become less dependent on grants and donations. When an institution realizes revenue generated from the sale of the SπC, less the fees and commissions, then the revenue can be invested in capital programs (e.g. social, environmental, built, or human capital investments). There is pressure on the organization to perform since failure to perform impacts future issuances. As a result, the particular sector under considerations will benefit from acceptances of best practices as well as merger and acquisition activity in this competitive environment. Those organizations that opted not to participate early in the market will have time to improve their performance in order to get a greater competitive return on their SπCs. However, we imagine that those organizations that are unable to improve their performance or are unwilling to participate in the SπC market may
cease to exist, thereby opening the playing field for organizations that can create $S\pi$ at higher value. The intent, and believed overall effect, however, should be to bring the best parts of market efficiency to a sector that has not typically had these benefits.

The creation of the third party Agent offers benefits to all parties involved in the $S\pi$C. Naturally, the Agent, if the job is properly done, benefits through the receipt of fees. For both the investor and the $S\pi$ firm, the Agent and the trade in $S\pi$C serve to make the activities of $S\pi$ firms more visible; the market in $S\pi$C would be more information-rich than the current situation. $S\pi$ firms will benefit from the informational scans that an Agent must perform, and the investor makes more informed choices about where and how to invest. At present, social profit enterprises are funded principally through government, socially-conscious private individuals, grant-funding institutions, and corporate entities seeking (usually non-controversial) tax deductions.

The $S\pi$C provides an additional funding mechanism, which maintains the usefulness of private and corporate donations, and offers monies that can be used in an unrestricted manner. Individuals, corporations, and grant-funding institutions frequently have demands upon the recipients of their funding, and such demands can limit a $S\pi$ firm’s ability to expand into new areas of social action, improve or develop human resources, or begin to think more critically and creatively about the $S\pi$ institutional mission. The $S\pi$C then offers an opportunity for innovation that is lacking in the current nonprofit world.

**BENEFITS FOR INVESTORS**

Perhaps the most important part of an $S\pi$C is having somebody willing to invest. Unlike the fixed and known return of the tax donation from philanthropy and giving, these risk takers must realize and accept that the return is not guaranteed, which means they may lose their entire investment (Callanan et al., 2012). While the primary market for investors in $S\pi$Cs may be “limited to philanthropic and socially minded investors willing to subsidize the achievement of social goals” (Liebman, 2011), we imagine that others will follow once performance is demonstrated. In theory, these “socially minded investors” are willing to accept a smaller return on the social project compared other possible investment options because they want to do what is best for society but with a higher potential than traditional tax offsets.

The $S\pi$C market increases the desirability of socially conscious corporate investment because of increased transparency, and greater $S\pi$-achievement. At present, socially conscious corporate investing is, for the most part, only a part of corporate interest in maximizing all opportunities to increase the ‘bottom line’ (Berber et al., 2011). However, investors that are looking for stability in return may be more attracted to this market since the overall risk is expected to be lower and since they realize the intrinsic value of the investment. Even if the nonprofit
venture fails, the investor can still write off the investment capital as a loss (thus decreasing taxable income) while contributing social good to society even if it did not measure up with the desired outcome (Dagher, 2013).

Nonetheless, the private investors in SπC contracts have the most to lose. They not only rely on the nonprofit organization to perform well but also on the Agent to pay when performance measures are met (Vogel & Klissurski, 2013). For this reason, it is particularly important to the investor that the recipient of their donation is capable of creating substantial social profit. It is up to the social profit enterprise to convince the investor that their plan of action will work and meet performance targets set in the bond. Because the investors rely on the Agent to provide an end-of-year issuance of tax credit to both government and investor, they have to also subscribe to the idea that this third party will credibly honor the agreement set forth in the instrument. One opportunity for government-involvement in this market would perhaps be initially in offering some guarantee that these Agents provide the necessary expertise and financial backing, much as they do in the banking industry, but, presumably, not with the same level of taxpayer risk (Sheffield, 2013).

CONCLUSIONS

It is clear to see that the creation of the credit and the marketplace has a long way to go. There is a lot of interdependence amongst the key entities, which cooperate to create SπCs: nonprofits, Agents, investors, and government. Nonetheless, there are analogues, such as the case in carbon markets in Europe. These analogues provide the foundation upon which to build the SπC market and demonstrate proof of concept. This newest proposal for a social impact bond relating to social profit organizations as a funding method is much different than the existing framework. The SπCs are sold specifically to provide unrestricted funding to the social profit organization. The credits are pure philanthropic mechanisms, and unrestricted funds are, in the end, the type of philanthropy most useful to social profit firms, and, therefore, stand to increase society’s overall social benefit. It is likely that a certain amount of resistance to this idea will appear. Certain criticisms of the idea of social impact bonds which has provided a basis for our concept have already emerged (MacDonald, 2013).

However, there are many potentially positive outcomes of a market for SπCs. First, there is an incentive for nonprofit organizations to be innovative and entrepreneurial. As new models are presented, new delivery mechanisms created, or new forms of products to meet the markets needs are introduced, innovators and entrepreneurs are rewarded for their creativity. Second, there is a focus on outcomes rather than on inputs that drives performance. Third, instead of blindly making donations in order to get tax write-offs, investment into SπCs becomes a source of competition, potentially both for the FP donating entities as well as for Sπ firms. Institutional and corporate investors are likely to conduct more research into their investments before making a choice in order to
ensure that they have an increased likelihood of success.

In this paper, we have proposed an outline of the SπC, a mechanism that can serve to improve performance in the nonprofit sector and thus increase overall social well-being. In order for the social profit credit to take effect further work is necessary. The tax codes would need to put the appropriate incentives in place. Mechanisms for evaluating performance and in various sectors need development; however, in some sectors, nascent indexes may already exist that can provide a basis for the third party SπC valuations. Finally, markets need to be created where the credits can be traded. Our proposal also will no doubt benefit from input and critique from other academicians and professionals currently working in nonprofit management. There are important areas that we may not have fully developed herein. For example, we make an assumption commonly critiqued by the nonprofit sector that market dynamics will work when applied to the nonprofit arena.

While it is true that we believe that nonprofit organizations may benefit from the innovation and competition that market dynamics encourages, there may be unintended consequences that need additional exploration. In addition, we have not fully addressed the impact that secondary markets for the social-profit credits may have on the initial offerings. Further, while the functional structure of the social-profit credit was developed within this work, additional efforts need to be conducted to fully structure the market for the credits as well as the tax and financial regulations necessary to implement these.

There appears to exist many interesting paths for future research within the management literature. For instance, while the SπC will clearly incentive competition in the nonprofit sector, it may also provide a source of competitive advantage for the commercial firm. Proper investment in SπC, like any investment mechanism, may result in abnormal gains for the firm when properly researched; however, it may also have the added benefit of continuing to provide increases to goodwill on the firm’s balance sheet. This line of inquiry can further be extended to included portfolio optimization of philanthropic giving be for-profit firms.

We conjecture that firms will look to better quantified results from nonprofits as they look to optimize their philanthropic investments. Finally, while we suggested in the body of this paper that the average performance of nonprofits will continue to rise as the lower performers are acquired, die, or adapt, testing of this proposition needs to be fully developed and explored.

REFERENCES


bonds 101.” *Center for American Progress*, pp.1–4.


ROTH IRAS AND QUALIFIED CHARITABLE DISTRIBUTIONS

Sheldon R. Smith  
Lynn R. Smith  
Utah Valley University

ABSTRACT

Current tax law allows certain taxpayers to make direct charitable contributions from an individual retirement account without including the transferred amount as income and then claiming it as a charitable contribution. This paper discusses this tax law in relation to a Roth individual retirement account (Roth IRA). While some distributions from a Roth IRA technically could qualify for this tax treatment, it would be very unusual for a distribution from a Roth IRA to meet all the qualifications for a qualified charitable distribution (QCD). Therefore, even though saying a QCD could come from a Roth IRA is technically correct, it is not very relevant from a practical standpoint.

Key Words: Roth IRAs, Qualified Charitable Distributions, Tax Law

INTRODUCTION

Internal Revenue Code (IRC) Section 408(d)(8) allows certain taxpayers to make direct charitable contributions from some individual retirement plans without including the transferred amount as income and then claiming it as a charitable contribution (IRC, Sec. 408). This tax law has several possible advantages to those who qualify to use it. It can also be beneficial to charities, as they may be able to get additional charitable contributions from taxpayers who can use this method to make charitable contributions. Roth IRAs technically qualify as plans that could be used for a qualified charitable distribution (QCD). However, it would be a very unusual case where a Roth IRA distribution could actually qualify for QCD treatment. This paper will explain how a QCD works and why a Roth IRA is seldom the appropriate vehicle to take advantage of this tax break.

Following this introduction, this paper will include sections on the historical background of qualified charitable distributions (QCDs), qualifications for QCDs, and possible advantages of QCDs. The paper will then introduce Roth IRAs and provide some background on how contributions and distributions work. The last two sections of the paper will provide a discussion of Roth IRAs and QCDs and a conclusion.
HISTORICAL BACKGROUND OF QCDs

IRC Section 408 relates to individual retirement accounts, and Section 408(d) discusses the tax treatment of distributions from such accounts. Section 408(d)(8) was created by Public Law 109-280 (2006), enacted on August 17, 2006. This law added the provision for a QCD, but it also included a termination date—December 31, 2007. Therefore, the law lapsed at the beginning of 2008.

However, as is often the case, tax laws with expiration dates included are often extended, but the extensions aren't necessarily legislated before the provision lapses. Public Law 110-343 (2008), signed on October 3, 2008, extended this tax code provision to the end of 2009 and made it retroactive to the beginning of 2008. The QCD again lapsed at the beginning of 2010 but was extended to the end of 2011 by Public Law 111-312 (2010). This law also made the provision retroactive to the beginning of 2010. However, because the extension was legislated so close to the end of 2010 without adequate time for taxpayers to make QCDs for 2010, this extension also allowed QCDs made before February 1, 2011 to be deemed to have been made on December 31, 2010, thus allowing them to qualify as QCDs for 2010.

True to form, the provision again lapsed at the beginning of 2012, only to be extended to the end of 2013 and made retroactive to the beginning of 2012 by Public Law 112-240 (2013). Because this extension wasn't actually signed until January 2, 2013, two special rules were included with the extension. The first was similar to the special rule from the prior legislation: QCDs made before February 1, 2013 were deemed to have been made on December 31, 2012, thus allowing them to count as QCDs for 2012. The second special rule allowed taxpayers who had received a cash distribution from an IRA after November 30, 2012 but before the end of that year to transfer the cash amount to a qualified charitable entity by February 1, 2013 and have it count as a direct transfer to the charity. QCDs must normally be directly transferred from the IRA trustee to the charitable entity, but this special rule allowed an intermediate transfer to the IRA owner to count as a direct transfer so it could qualify as a QCD.

The provision again lapsed at the beginning of 2014 but was extended to the end of 2014 with retroactive treatment to the beginning of 2014 by Public Law 113-295. Even though this extension was legislated late in 2014, no special rules were included with this update, so in this case, besides being only a one-year extension, it did not allow taxpayers much time to make their QCDs from their IRAs by the end of the 2014 tax year. One more time the provision lapsed at the beginning of 2015. However, near the end of 2015, additional legislation in Public Law 114-113 (2015) made this tax provision retroactive to the beginning of 2015 and eliminated the termination date for Section 408(d)(8). So while no tax law can be considered permanent, the specific law allowing for QCDs will exist unless specific legislation eliminates it. With the elimination of a termination date for this tax law, taxpayers can make better long-term plans how to use this tax law to
their advantage. Charities will also be better able to use this tax law to encourage certain donors to make more and/or larger contributions over time since the uncertainty of lapses and extensions is now gone.

**QUALIFICATIONS FOR QCDs**

To count as a QCD, certain conditions must be met. The taxpayer must be at least 70½ years old at the time of the contribution. The contribution must be made directly from the IRA to the charity. Distributions from a simplified employee pension (SEP, sometimes referred to as a SEP-IRA) or a simple retirement account (sometimes referred to as a SIMPLE IRA—Savings Incentive Match Plan for Employees) cannot qualify as QCDs. Most charitable organizations can receive these distributions, but contributions to private foundations or to donor-advised funds cannot be QCDs. The annual limit for a QCD is $100,000, but this is an individual limit, so if spouses both contribute $100,000 from separate IRAs during the year, joint filers could contribute up to $200,000 annually through the use of a QCD (IRC, Sec. 408).

Of course, the taxpayer will not be able to take a charitable contribution deduction for the amount contributed through a QCD, as the tax benefit is already received by excluding from income the amount that would normally be a taxable IRA distribution. In addition, the full amount of the distribution must otherwise qualify as a charitable deduction to be a QCD. Charitable contributions can be deducted as itemized deductions to the extent they exceed the value of any benefit the donor receives from the charity (IRS, 2016b). For a transfer from an IRA to qualify as a QCD, the entire amount contributed must otherwise qualify as a charitable contribution; the taxpayer cannot receive any benefit from the organization for the contribution (IRC, Sec. 408). Another requirement which must be met for the taxpayer to claim a distribution as a QCD is that the taxpayer must receive an appropriate acknowledgement from the charity to document the contribution (IRS, 2016d).

A distribution from an IRA can only be a QCD to the extent that it would otherwise be taxable without this specific tax provision. Many IRAs only have pre-tax contributions and earnings; all distributions from these IRAs would normally be taxable. However, all distributions from these IRAs that meet the other QCD requirements would be excluded from income. On the other hand, some IRAs include nontaxable contributions (made on an after-tax basis) that were not deductible at the time of the contribution, as well as taxable earnings and contributions. For regular distributions from such an IRA, a portion of each distribution is taxable and a portion is nontaxable (until all of the nontaxable portions have been withdrawn) (IRS, 2016d). In determining the taxable and nontaxable portions of distributions, all IRAs are treated as one contract and all distributions for a year are treated as one distribution (IRC, Sec. 408).
To make the QCD provision more beneficial, a special ordering of IRA distributions is allowed. This special ordering allows direct distributions to charities to come from the portions of the IRAs which are taxable first, again treating all IRA contracts as one. Thus, a taxpayer with an IRA which has nontaxable portions can consider distributions to charities to come entirely from the taxable portion up to the extent of any taxable portion in the IRAs (IRC, Sec. 408). Of course, recalculations would need to be made after any QCD to determine taxable and nontaxable portions of any future IRA distributions.

**ADVANTAGES OF QCDS**

Several possible advantages of QCDs for those who qualify have been enumerated. These different advantages may apply to varying groups of taxpayers who otherwise qualify for the QCD. One tax advantage is that those who are not itemizing deductions on their tax returns can still take the standard deduction but get a tax advantage of contributing to charity by not having to claim the IRA transfer as income (Independent Sector, 2016). A second potential advantage relates to the percentage limits on charitable contributions. While charitable contributions have limits, typically 50 percent of the contribution base (usually adjusted gross income (AGI) for individual taxpayers) for cash contributions to most charities, the contribution through a QCD avoids these limits, allowing taxpayers to contribute more to charity (Smucker, 2016).

Another advantage ties to the taxability of social security benefits. Up to 85 percent of social security benefits can be subject to federal income taxes, depending on income level. If charitable contributions are made directly through a QCD, the IRA distribution is not counted as income and will not increase the taxability of any social security benefits received (American Endowment Foundation, 2016). Tax law provides for the phaseout of itemized deductions and personal exemptions as income increases. Making a charitable contribution through a QCD will not cause the taxpayer to enter or move further into the phaseout range, perhaps saving the taxpayer some amounts of these benefits that might otherwise be lost if the IRA distribution were made to the taxpayer who then contributed the amount to charity (Smucker, 2016).

Miscellaneous itemized deductions can only benefit a taxpayer to the extent they exceed a specific floor, typically 2 percent of AGI. Medical expenses can only be itemized to the extent they exceed 10 percent of AGI (or 7.5 percent of AGI for those over 65 through 2016). By using a QCD for a charitable contribution, the taxpayer can avoid raising the floor for these types of itemized deductions, potentially allowing a greater benefit from these deductions (Smucker, 2016).

State taxes may also be reduced through a QCD, especially for taxpayers in states that require taxpayers to begin with the federal AGI on the state tax return. Some states do not allow a deduction for charitable contributions. The QCD might
allow taxpayers in these states to get a tax benefit from this type of charitable contribution even though an IRA withdrawal followed by a charitable contribution would not provide a state tax advantage (American Endowment Foundation, 2016).

Public Law 111-152 (2010) created a new tax effective starting in 2013, a 3.8% net investment income tax (NIIT) on the smaller of (1) net investment income, or (2) the excess of modified adjusted gross income over a threshold amount (depending on filing status). Someone who otherwise qualifies for a QCD and who has net investment income might be able to reduce or eliminate the NIIT due by making a QCD from an IRA distribution that would otherwise be taxable and be subject to the NIIT.

By the time a taxpayer reaches age 70½, distributions from a traditional IRA are required. These are called required minimum distributions (RMDs), with the amounts recalculated each year based on account balances and expected lifespan. RMDs are required so the government can start to collect taxes on these amounts that have grown tax free, including the amounts originally contributed which, in some cases, were pre-tax amounts. Any amount contributed through a QCD counts toward that annual RMD but is not part of taxable income. In addition to excluding this amount from income, it might also keep a taxpayer from moving into a higher tax bracket, thus reducing the tax liability even further (Smucker, 2016). The American Endowment Foundation (2016) also mentions future estate tax implications. Since the estate and income taxes on IRAs can be significant, a QCD allows a taxpayer to reduce the estate by contributing to charity through a QCD without federal income tax consequences.

ROTH IRAS

Traditional IRAs were first allowed by Public Law 93-406 (1974). Laws relating to traditional IRAs can be found in IRC Section 408. Roth IRAs were first allowed over 20 years later when Public Law 105-34 (1997) added IRC Section 408A. While traditional IRAs offer certain taxpayers deductibility for their contributions to the accounts, Roth IRAs offer no deductibility; all contributions to Roth IRAs are made with after-tax funds.

Roth IRAs can be opened at any time. In addition, no age limit exists for creating a Roth IRA or for contributions to a Roth IRA, as exist for a traditional IRA. Regular contributions for a specific year can be made to a Roth IRA at any time up to the original due date of the tax return for that year (IRS, 2016c); subject to an annual dollar limit, a limit based on compensation, and a limit based on modified AGI. The annual limit for 2017 is $5,500 (or $6,500 for those over age 50) (IRS, 2016a). This limit is for all IRA contributions made by the individual, so if someone contributes to a traditional IRA, any amount contributed there would reduce the amount that can be contributed to a Roth IRA. The amount contributed cannot exceed the taxable compensation of the individual for that year. An
exception exists for married couples filing jointly; if one member of the couple has enough taxable compensation to cover both individuals’ contributions, both can make contributions even if the other individual did not have enough taxable compensation (IRS, 2016c). The ability to contribute to a Roth IRA is also limited by modified AGI. For 2017 a couple filing jointly can contribute the full dollar limit if the modified AGI is less than $186,000. However, for modified AGI between $186,000 and $196,000, the amount that can be contributed is reduced. If the modified AGI is $196,000 or more, no contribution can be made to a Roth IRA for that year (IRS, 2016a).

Besides regular contributions to a Roth IRA, an individual may rollover other retirement plans (such as a 401(k) plan, a 403(b) plan, or a 457 plan) into a Roth IRA or convert amounts from other IRAs (such as traditional IRAs, SEP-IRAs, and SIMPLE IRAs) to Roth IRAs. However, any amounts rolled over or converted from other plans which had not yet been taxed would become taxable at the time of the rollover or conversion (IRS, 2016c).

The simple idea behind a Roth IRA is that after-tax contributions are made to the account, but all Roth IRA distributions, including earnings, can be withdrawn totally tax free when certain conditions are met. Distributions from a Roth IRA that represent regular contributions are not taxable, as the regular contributions were taxed before being contributed. In addition, qualified distributions are not taxable. A qualified distribution must meet two requirements. The first requirement can be met in multiple ways: (1) the account holder receiving the distribution is at least 59½, (2) the distribution is made to a beneficiary or estate upon the death of the account holder, (3) the distribution is made because the individual is disabled, or (4) the distribution meets specific rules for a first-time homebuyer. The second requirement is that the distribution must come after the fifth taxable year for which the individual first made a contribution to a Roth IRA (IRS, 2016d).

If a Roth IRA distribution is not a qualified distribution, part of the distribution may be taxable. To determine the taxable amount, a specific ordering is used. Regular contributions are considered to be distributed first. Rollover and conversion contributions are considered to be distributed next, with the taxable portion (taxable at the time of the rollover or conversion) taken before the nontaxable portion. Earnings on contributions are considered distributed last. This provides maximum benefit to the individual, as earnings, which would be the taxable portion of a nonqualified distribution, are not claimed until all contributions are distributed (IRS, 2016d).

If distributions from a Roth IRA are not qualified distributions, they are termed early distributions and may be subject to an additional tax of 10 percent (as a penalty) on any taxable portion of the distribution. This tax is in addition to the regular income tax that would be assessed. Early distributions from a rollover or
conversion are also subject to this 10 percent additional tax for the five-tax-year period from when the rollover or conversion was made. This five-year rule applies separately to each rollover or conversion into a Roth IRA account. Any distribution of these amounts before five tax years have passed makes the original amount which was taxable at the time of the rollover or conversion subject to an additional 10 percent tax because it is treated as an early withdrawal. The additional 10 percent tax would also apply to any taxable earnings on these contributions. However, a number of exceptions exist which would eliminate the 10 percent penalty tax, including if the individual is at least 59½ (IRS, 2016d).

Because all distributions from a Roth IRA can be tax free when the requirements are met, the government has no incentive to require an account holder to take required distributions. Therefore, unlike a traditional IRA, a Roth IRA has no RMD rules. This is true as long as the original account owner is alive (IRC, Sec. 408A). For inherited IRAs, different options exist, depending on who is the beneficiary. If the spouse is the sole beneficiary, more options exist than for other beneficiaries. Inherited Roth IRAs may have RMDs, but if the account has been open for five years, distributions will be qualified distributions and will not be taxable. In addition, if a beneficiary receives distributions from a Roth IRA of a decedent, the 10 percent additional tax will not apply. However, if distributions are received from inherited Roth IRAs which have not met the five-year rule, the earnings will be taxable (IRS, 2016d; Charles Schwab, 2016).

DISCUSSION

Sources indicate that QCDs must be contributions from traditional or Roth IRAs (Joint Committee on Taxation, 2013; Topoleski, 2014). Several definitions from the tax code are relevant in determining how a distribution from a Roth IRA could qualify as a QCD. IRC Section 408A(a) states that “except as provided in this section, a Roth IRA shall be treated for purposes of this title in the same manner as an individual retirement plan.” IRC Section 408A(b) indicates that “the term ‘Roth IRA’ means an individual retirement plan (as defined in section 7701(a)(37)) which is designated . . . at the time of establishment of the plan as a Roth IRA.” IRC Section 7701(a)(37) defines what is meant by an individual retirement plan. “The term ‘individual retirement plan’ means—(A) an individual retirement account described in section 408(a), and (B) an individual retirement annuity described in section 408(b). These references to Section 408 refer to the code section on individual retirement accounts. Since a Roth IRA is specifically defined as an individual retirement plan, the rules in Section 408 which apply to an individual retirement plan would also apply to Roth IRAs in Section 408A unless Section 408A specifically states otherwise. For example, IRC Section 408(d)(1) indicates that “except as otherwise provided in this subsection, any amount paid or distributed out of an individual retirement plan shall be included in gross income by the payee or distribute. . . .” However, IRC Section 408A(d)(1) counters that mandate when it states that “any qualified distribution from a Roth IRA shall not be includible in gross income.”
By definition, a transfer from an IRA can only be excluded from income as a QCD if it is made by someone who is at least 70½ and if it would otherwise have been a taxable distribution from the IRA. But if a Roth account holder is at least 59½ and made a contribution to any Roth IRA at least five tax years previously, all Roth IRA distributions for that individual would be excluded from gross income as qualified distributions. Since any Roth IRA account holder who is at least 70½ is also 59½, if this individual has had a Roth IRA for at least five tax years, no distribution from the Roth IRA could be taxable; therefore, no distribution could qualify as a QCD.

As mentioned earlier, for Roth IRAs, distributions are considered to come first from regular contributions, then from rollover and conversion contributions, and then from any earnings. This would make taxable earnings, if any, come out last (IRS, 2016d). However, the special ordering rule for both traditional and Roth IRAs means that distributions used for a QCD are assumed to come first from taxable portions of the traditional or Roth IRA, aggregating the accounts in each category and assuming all amounts available were distributed during the year (IRC, Sec. 408(d)(8)(D)). This allows the maximum possible advantage of a QCD since taxable amounts come out first and any amount used for a QCD must be taxable.

A Roth IRA account holder, although at least 70½, who has not met the five-year rule for having the account open, could have some earnings in the account which would be taxable if distributed, as the earnings would not be qualified distributions nor would they be contributions which had already been taxed. If these earnings were distributed to the account holder, they would be subject to income tax but would not be subject to the additional 10 percent penalty tax, as someone over 59½ is exempt from that additional tax. In this case, it is possible for the taxable earnings to be contributed as a QCD. If so, these earnings would not be taxable and might help the taxpayer achieve one or more of the possible benefits of a QCD mentioned earlier.

For inherited Roth IRAs, if the account has been open for five years, no portion of the distributions would be taxable; thus, no portion could be used for a QCD. But if a Roth IRA has not been open for five years and is then inherited, the earnings from distributions could be taxable and therefore qualify for QCD treatment. Since most taxpayers over age 70½ who have Roth IRAs would have had them open for more than five years, it would be very unusual for a distribution from a Roth IRA or an inherited Roth IRA to be a QCD. Although this can happen, it would be unlikely. In addition, the beneficiary of a Roth IRA may not be 70½, in which case the QCD rules would not have been met. Therefore, although the IRC technically allows QCDs to come from a Roth IRA, it is probably not very relevant for taxpayers or their financial advisors to consider a QCD from a Roth IRA.
IRA. It would also not be very helpful for charities to tout this method of giving from any IRA except a traditional IRA.

CONCLUSION

U.S. tax law allows for direct transfers of otherwise taxable income from IRAs to charities to be excluded from income if all of the rules for a QCD are met. This law has several possible tax benefits for different groups of older taxpayers who have IRAs and want to make contributions to charity. It can also benefit charities which may be able to increase the amount and/or number of contributions because of this tax law. Because the termination of this tax law has now been removed, taxpayers and charities will be better able to make long-term plans for using or encouraging the use of this tax provision.

Although a direct transfer from a Roth IRA to a charity might technically qualify for at least partial QCD status, the rules that apply to QCDs mean that most distributions from Roth IRAs would not qualify for the tax exclusion because they would not be otherwise taxable. While a taxpayer could make a direct transfer from a Roth IRA, the nontaxable portion would not qualify for QCD status but could qualify as a charitable contribution which could be deducted as part of a taxpayer’s itemized deductions. Since it is nontaxable anyway, there would be no specific advantage of using a direct transfer from the account to the charity, as the same thing could be accomplished by taking a cash distribution from the IRA and then using the cash to make a charitable contribution which could be deducted if the taxpayer itemizes deductions. Even if a portion of the distribution qualified because it came from the earnings in an account that had not yet been opened for five tax years, it would likely be a relatively small amount compared to the balance in the Roth IRA. So it is likely that most taxpayers will want to work only with traditional IRAs as they consider the tax advantages of a QCD as a way to make charitable contributions.

REFERENCES


Joint Committee on Taxation. (2013). Present Law and Background Relating to the Federal Tax Treatment of Charitable Contributions (JCX-4-13), February 11.