PARKER
COLLEGE OF BUSINESS
GEORGIA SOUTHERN UNIVERSITY

Department of Economics
11935 Abercorn Street
Savannah, GA 31419

Farmingdale State College

State University of New York

School of Business
2350 Broadhollow Road
Farmingdale, NY 11735-1021

ACADEMY of
ECONOMICS
and FINANCE

PAPERS AND PROCEEDINGS

56th
ANNUAL MEETING

ST. PETE BEACH, FLORIDA
FEBRUARY 6-9, 2019

editors
DR. MICHAEL TOMA
DR. RICHARD VOGEL

Academy of Economics and Finance

Papers
and
Proceedings

Volume 43, 2019

Fifty-sixth Annual Meeting
St. Petersburg Beach, FL
February 6-9, 2019

Program Arranged by Olivier Maisondieu Laforge
University of Nebraska Omaha

Editors

Michael Toma, Georgia Southern University
Richard Vogel, Farmingdale State College, State University of New York

2018-2019 AEF Leadership

President, Michael Toma, Georgia Southern University

President Elect, Olivier Maisondieu Laforge, University of Nebraska Omaha

First Vice President, Speros Margetis, University of Tampa

Second Vice President, Mary Funck, Sam Houston State University

Secretary, Kath Henebry, University of Nebraska Omaha

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

Local Arrangements Chair, Speros Margetis, University of Tampa

Directors

Stephen Conroy, University of San Diego, CA (2019)

Abeba Mussa, State University of New York at Farmingdale, NY (2019)

Marc Sardy, Rollins College, FL (2020)

Adam Jones, University of North Carolina, Wilmington, NC (2020)

Robert Burney, Coastal Carolina University (SC) (2021)

Michael McLain, Hampton University (NC) (2021)

M. Kabir Hassan, University of New Orleans, (LS) (2022)

James Malm, College of Charleston (SC) (2022)

Recent Presidents

Michael Toma, Georgia Southern University

Shadhdad Naghshpour, Alabama A&M University

Edward J. Graham, University of North Carolina Wilmington

M. Douglas Berg, Sam Houston State University

Jocelyn Evans, College of Charleston

Robert T. Burrus Jr., University of North Carolina Wilmington

Steven Jones, Samford University

Robert Stretcher, Sam Houston State University

Richard McGrath, Armstrong Atlantic State University

Bill Sackley, University of North Carolina Wilmington

David Boldt, University of West Georgia

Albert DePrince, Jr., Middle Tennessee State University

Editors, AEF Papers and Proceedings
Michael Toma, Georgia Southern University
Richard Vogel, Farmingdale State College

Editor, Journal of Economics and Finance
James Payne, University of Texas El Paso

Editor, Journal of Economics and Finance Education
E. F. Stephenson, Berry College

Editors, AEF Journal
Clay M. Moffett, University of North Carolina Wilmington
Michael Toma, Georgia Southern University

Webmaster
Doug Berg -- Sam Houston State University

Academy of Economics and Finance

Volume 43	**Papers and Proceedings**	**2019**

Table of Contents

Predictors of 6-Year Graduation Rates Among African-American Students At Selected Public Universities: Do HBCUs Make A Difference?

Richard F. Bieker, Delaware State University

Abstract

A major issue faced by institutions of higher education in the U.S. is the very low 6-year graduation rates of African American students. The objective of this study is to examine whether the 6-year graduation rates of African American students differ significantly between public Historically Black Colleges and Universities (HBCUs) and similar Historically White Colleges and Universities (HWCUs) after controlling for other differences between the two types of institutions. The 6-year graduation rate is used to compare the performance of the two types of institutions with respect to educating African American students because it is the most widely used metric of college and university performance. Data used in the analysis were obtained from the Integrated Postsecondary Education Data Systems (IPEDS).

Introduction

Today's Historically Black Colleges and Universities (HBCUs) have their origins in the post-Civil War period in the southern and border states. In the beginning of this period there was de facto segregation of the races in higher education. The Morrill Act of 1890 established and provided funding for what are known today as the 1890 land-grant institutions (Montgomery & Montgomery, 2012). The Plessy v. Ferguson Supreme Court decision of 1896 formalized the segregated two-tier system of public land grant institutions in a legal (de jure) sense (Fryer & Greenstone, 2007). There are currently 19 HBCU land grant institutions in 18 states. In addition to these land-grant institutions, other Black colleges were established by religious and benevolent organizations and by some states in the post-Civil War period (Kim, 2002). And, prior to the middle of the twentieth century, more than 90% of all African American students who were enrolled in institutions of higher education were enrolled at HBCUs (Kim and Conrad, 2006).

This two-tiered system of higher education in the South continued until the Brown v. Board of Education Supreme Court Decision in 1954 and subsequent legislation. Because of the desegregation that was brought about by the Brown v. Board of Education decision, rising incomes among African American households, and easier access to financial aid, African American enrollment in institutions of higher education increased significantly beginning in the early 1960s. However, during the same period, the proportion of African American college students enrolled at HBCUs began to rapidly decline. By the fall of 1980 only 17 percent of African American college students were enrolled at HBCUs. That share declined further to a mere nine percent by 2015 (Anderson, 2015).

At the same time that this desegregation was occurring, and a greater proportion of African American students were matriculating at Historically White Colleges and Universities (HWCUs), the level of public funding for public HBCUs increased significantly. Title III of the Higher Education Act of 1965 was enacted to provide funding for HBCUs "to strengthen academic quality, institutional management, and financial stability" (Mercer, 2008). In addition, while the main intent of the 1973 U.S. District Court for the District of Columbia decision in the Adams v. Richardson case was to desegregate higher education, it also required states to increase funding for HBCUs. This occurred because the Court ruled that states could not satisfy their desegregation compliance requirements by closing HBCUs. Instead, states were directed to increase their financial support for HBCUs (Egerton, 1975; Fryer & Greenstone, 2007).

While some scholars have been focused on the role that HBCUs play in providing higher education opportunities for African Americans, a wider concern about the cost and quality of higher education in the U.S. began to emerge during the last decade of the twentieth century. And, this concern has since become more intense. Some of the long-held assumptions about the quality and value of higher education in the U.S. began to be subjected to increased levels of scrutiny. The issuance of the report entitled *A Test of Leadership: Charting the Future of U.S. Higher Education* by the commission appointed by Secretary of Education Margaret Spellings in 2006 concluded that higher education in the U.S. was found lacking on several fronts. The commission was critical of institutions of higher education regarding cost and affordability, lack of transparency and accountability, lack of innovation, and lack of assurance of learning (U.S. Department of Education, 2006). Higher education came under further criticism with the release of the book entitled *Academically Adrift* in 2011. In this book, Arum and Roksa (2011) reported the results of tracking the performance of 2,300 college students. They measured the changes in the students' critical thinking and analytical skills over the period and found that 36 percent of the students showed no improvement in

learning over their four-year period of matriculation and that those students who did show improvement showed only marginal improvement. At the same time that questions about the quality and value of higher education were being raised, there began to be increasing concerns about the rising costs of higher education and the level of debt that students and their families were incurring to attend college. By the second quarter of 2018 total outstanding student loans reached $1.53 trillion (Singletary, 2018).

Given these heightened concerns about the cost and quality of higher education in the U.S. in recent years, there have been increasing calls for colleges and universities to be held more accountable to their stakeholders (Archibald and Feldman, 2008; Morrison, 2015; Slaughter, 2017; Smith, 2018). Within this broader context, one question that is being asked is whether there continues to be a compelling justification for the continued financial support of public HBCUs. That is, do HBCUs have a unique and significant purpose that would justify their continued public funding (Gasman et al., 2007; Gasman & Bowman, 2011; Riley, 2010; Stewart, 2011)? There have been a number of studies that have attempted to address this issue. However, the results of these studies are inconclusive and in some cases contradictory. In addition, the studies are now quite old. The purpose of this study is to attempt to provide further insight into the question.

This study differs from previous studies in that it focuses on comparing African American 6-year graduation rates at public HBCUs with those at comparable public HWCUs that are located in the same states. Specifically, the objective of this study is to determine whether the 6-year graduation rates of African American students are significantly different between HBCUs and their peer public HWCUs after controlling for other differences between the two types of institutions. This is an important research question for several reasons. First, the public HWCUs in the sample, except for differences in the demographic composition of the student body, are similar to the sample public HBCUs in terms of mission, admission standards and sources of funding. In addition, it could be argued that these are the HWCUs that the African American students who are currently enrolled in the sample HBCUs would most likely be admitted to and attend if they did not have access to HBCUs. This is because the two types of institutions are geographically close to one another and have similar missions, admissions standards and tuition and fees. In fact, there have been proposals that some of the public HBCUs be merged with HWCUs that are included in the sample utilized by this study (Stewart, 2011).

The 6-year graduation rate is used to compare the two types of institutions because it is the most widely used metric of college and university performance (Archibald and Feldman, 2008; Anstine, 2013; Campbell and Fuqua, 2008; Montgomery and Montgomery, 2012). This is likely the case because the federal government's Federal Student Right-to-Know and Campus Security Act of 1991 requires that colleges and universities publish data on graduation rates. In addition, the 6-year graduation rate is the measure of performance of higher education institutions that is most frequently used by state governments to evaluate these institutions (Archibald and Feldman, 2008). A study of the relative effectiveness of the two types of institutions in generating high 6-year graduation rates for African American students is particularly important because, nationally, recent 6-year graduation rates of African American students are significantly below those of white students. For example, white students who started as full-time students in 2009 had a 6-year graduation rate of 58.7 percent compared to only 36.7 percent for African American students who started as full-time students at the same time (Ginder, Kelly-Reid & Mann, 2016).

Literature Review

Early studies on the relative effectiveness of HBCUs in the education of African American students tended to be based on case studies or very small samples. Based on his review of early data and research, Allen (1987, pp.30) generalized that on black campuses, black students purchase psychological well-being, cultural affinity, nurturing academic relations, and happiness at the cost of limited physical facilities, fewer resources, and more restricted academic programs. On white campuses, black students purchase impressive physical plants, rich resources, and more diverse program options at the cost of social alienation, intense academic competition, racial stress, and loss of peace of mind.

However, despite Allen's claims, Wenglinsky (1996) concluded that the results of these early studies were largely anecdotal and inconclusive.

Later, Allen (1992) collected survey data for 1,800 African American students who attended six HWCUs and eight HBCUs for the years 1981-1983. Based on his findings, Allen (1992, pp. 39) concluded that …students in the sample who attended historically Black universities reported better academic performance, greater social involvement, and higher occupational aspirations than Black students who attended predominantly White institutions. In short, the college experience was most successful (measured by these outcomes) for African-American students on campuses with Black majority student populations.

There are, however, some serious limitations in trying to use the findings of Allen's study to respond to the research question posed by this study. First, the HWCUs that Allen selected for inclusion in his analysis were large and prestigious PhD granting institutions, including the University of Michigan, University of North Carolina, UCLA, Arizona State University, University of Wisconsin and the State University of New York at Stony Brook. These institutions are not likely the HWCUs that African American students who are currently attending HBCUs would be admitted to or choose to attend if the option of attending an

HBCU were not available. Also, the study is based on data from nearly 40 years ago. It is likely that much has changed at both HBCUs and HWCUs since that time.

In a study using student level data for 1,429 students from the 1990 National Postsecondary Financial Aid Study, Wenglinsky (1996) found that African American students at HBCUs were more likely to pursue a postgraduate education and become professionals than African American students attending HWCUs. However, African American students who attended HBCUs were no more likely than those attending HWCUs to experience high levels of social interaction with faculty and other students. Wenglinsky's study is limited in that it does not control for the type of HWCU that the African American students attended. Nor does it include any institutional variables except for a variable to measure the degree of student-faculty interaction. Finally, the data used in the study are likely too old to be of much benefit in addressing the question posed by this paper.

In a study using a longitudinal data set sponsored by the Higher Education Research Institute at UCLA and containing data for 941 African American college students who were followed for the period 1985-1994, Kim and Conrad (2006) compared the degree completion rates for African American students who attended HBCUs with those who attended HWCUs. Degree completion was posited to be a function of student and institutional characteristics. The authors found that African American students who attended HWCUs came from more affluent households and were better prepared academically as measured by GPAs and SAT scores than those who attended HBCUs. In addition, the HWCUs which the sample students attended tended to have a higher percentage of faculty with a PhD degree, higher faculty salaries, and higher instructional expenditures per FTE student than the HBCUs that the sample students attended. The authors did not find any difference in the probability of degree completion between the two types of institutions after controlling for differences in the other institutional and individual characteristics between the two groups. However, their study does have some severe limitations for purposes of answering the research question posed by this study. First, the data set is not current. In addition, since the student is the unit of analysis, it is not clear which HWCUs were attended by the students or whether these institutions were public or private.

In a more recent study Montgomery and Montgomery (2012) used purposive sampling to select 10 HBCUs and 10 HWCUs to compare the 6-year graduation rates of HBCUs with those of matched HWCUs, and to examine whether differences in institutional characteristics affected graduation rates. They used data for the years 2000, 2002, 2004, 2008 and 2010 and found that the overall graduation rates at HWCUs were higher than those at HBCUs. However, African American graduation rates at the two types of institutions were "virtually equal." The usefulness of the Montgomery and Montgomery paper is limited for purposes of answering the research question posed by this study for a number of reasons. First, the authors did not test whether the differences in 6-year graduation rates were statistically significant. In addition, the only institutional characteristic that was included in the analysis, other than whether the institution was an HBCU or a HWCU, was whether the institution was public or private.

In summary, the research conducted up to this point does provide useful insights into how an analytical framework might be developed to evaluate the research question posed by this study. However, none of the studies respond specifically to the research question posed by this study. Specifically, the research question that is addressed by this study is as follows: Do the 6-year graduation rates of African American students differ significantly between pubic HBCUs and their peer public HWCUs after controlling for other differences between the two types of institutions?

The Conceptual Framework and Hypotheses

The conceptual framework for this study is couched in terms of the theory of college departure as articulated by Tinto (1993) and the production function framework as proposed by economists (Webber & Ehrenberg, 2010). These two conceptual frameworks are complementary rather than competitive.

Tinto (1993) posited the process from college entry to college exit as a longitudinal one. A student enters college with a set of attributes including aptitude, motivation, academic achievement and socioeconomic background. These attributes determine the student's goals and motivation at the time of entry into the university. These pre-entry attributes affect the student's ability to integrate into and benefit from the social and academic environment of the institution. The degree of integration in turn will affect the student's ability to benefit from interaction with the faculty and engage in extracurricular activities. These in turn affect the student's academic performance and the likelihood that the student will persist and graduate. At the same time, characteristics of the institution that the student attends such as selectivity, size, diversity, and resources will affect the institution's impact on student persistence.

Using the economist's concept of the production function as proposed by Webber and Ehrenberg (2010), the 6-year graduation rate of students (GR) at school in time t can be modeled as a function of institutional inputs X, institutional characteristics Y, and student characteristics Z. That is:

$$GR_{it} = f(X_{it}, Y_{it}, Z_{it}) \hspace{3cm} (1)$$

Given the Tinto and economic production function conceptual frameworks the following conceptual model is formulated to account for differences in 6-year graduation rates among African-American students at the study institutions:

$$6YRGR_i = f(HSGPA_i, SAT_i, PCTPELL_i, PCTFTFACi, ACADSUPT_i, UGRADFTE_i, HBCU_i) \quad (2)$$

Where:
$6YRGR_i$ is the percent of first-time, full-time bachelor's or equivalent degree-seeking African American students who graduate within 6 years at the ith institution.
$HSGPA_i$ is the average high school GPA among freshman students at the ith institution.
SAT_i is the median SAT or equivalent ACT score of students admitted to and enrolled at the ith institution.
$PCTPELL_i$ is the percentage of full-time degree seeking students who are receiving Pell Grants at the ith institution.
$PCTFTFACi$ is the number of full-time faculty as a percent of all faculty at the ith institution.
$ACADSUPT_i$ is the expenditures on student support services plus academic support (not including instructional expenditures) per FTE at the ith institution.
$UGRADFTE_i$ is the number of full-time undergraduates plus the number of part-time undergraduates divided by three.
$HBCU_i$ is a dummy variable that denotes whether the ith institution is an Historically Black College or University (HBCU=1, HWCU=0).

The variables HSGPA, and SAT are intended to measure how well-prepared students admitted to a university are to succeed academically at the university level and to eventually complete their degree requirements. It is posited that both of these variables are positively related to the 6-year graduation rate.

The variable PCTPELL is intended to capture the effect that the socioeconomic background of the students enrolled at an institution have on the institution's 6-year graduation rate. Since the Pell grants are needs based, whether or not a student receives a Pell grant is a good proxy for the student's socioeconomic status. In accordance with previous research (Winkle-Wagner & McCoy, 2018; Zarifa, Kim, Seward, & Walters, 2018), it is posited that the 6-year graduation rates vary inversely with PCTPELL.

The variables PCTFTFAC and ACADSUPT are designed to capture the institutional commitment to instruction. It is posited that the 6-year graduation rate will vary directly with the percent of full-time faculty and the level of expenditures on student and academic support. It is likely that additional institutional inputs such as physical facilities, laboratory equipment and computer technology also affect 6-year graduation rates. However, the database used in this study did not contain data for these variables.

The variable, UGRADFTE is included to attempt to estimate the effect of the size of the institution on the 6-year graduation rates of African American students. Previous research has found that there are economies of scale in higher education (Cohn, Rhine & Santos, 1989; Zhang & Worthington, 2018). This suggests that, up to a point, as the full-time undergraduate enrollment at an institution increases the likelihood that the institution will acquire the specialized resources (e.g. teaching-learning centers and tutorial services) that will serve to enhance retention and 6-year graduation rates. However, other research by Millea, Wills, Elder and Molina (2018) indicates that, on campuses with more students and large class sizes, students are less likely to receive individualized attention and may feel less connected to their academic experience. As a result, retention rates were found to decline. This problem may be particularly acute for African American students who matriculate at larger HWCUs. For these reasons the terms UGADFTE and $UGRADFTE^2$ are included in the equation. It is posited that the coefficient for the UGADFTE term will be positive and that for the $UGRADFTE^2$ will be negative.

Finally, the dummy variable HBCU is included in the equation to determine if, after controlling for the other factors posited to affect 6-year graduation rates, there is a difference between the graduation rates of HBCUs and HWCUs.

The Data and Finding

Data used in the analysis were obtained from U.S. Department of Education's Integrated Postsecondary Education Data System (U.S. Department of Education, 2018). The data are for public institutions that are classified as Bachelor and Master level institutions by the Carnegie Foundation's classification system and which are located in the eighteen states in which the 19 HBCU land grant institutions that were created by the Morrill Act of 1890 are located. The 18 states are Alabama, Arkansas, Delaware, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, Missouri, North Carolina, Ohio, Oklahoma, South Carolina, Tennessee, Texas, Virginia and West Virginia. Using these selection criteria yielded a sample of 117 institutions. Twenty-six of the institutions were HBCUs and 91 were HWCUs. The data are for the academic year 2015-2016.

The study was restricted to institutions that are classified as Bachelor and Master level institutions by the Carnegie Foundation because all of the HBCUs in the sample states consisted of these types of institutions. Hence, it was deemed appropriate to include only similarly classified HWCUs.

Summary statistics by type of institution are shown in Table 1. Differences in African American 6-year graduation rates were not statistically significant between the two types of institutions. However, all of the explanatory variables were significantly different at least at the .10 level.

Students enrolled at HWCUs had significantly higher GPAs and SAT scores. On the other hand, the percentage of students receiving PELL grants was significantly higher at HBCUs than at HWCUs.

Table 1: Descriptive Statistics by Type of Institution

Variable	All Schools n=117	HBCUs n=26	HWCUs n=91	t/z value	Significance Level
6YRGR	32%	30%	33%	.298[a]	.765
HSGPA	3.26	2.89	3.37	8.741[b]	.025
SAT	985	868	1,018	11.319[b]	.043
PCTPELL	44%	65%	38	7.215[a]	.000
PCTFTFAC	64%	73%	63%	3.337[a]	.001
ACADSUPT	$2,393	$1,668	$2,600	2.469[b]	.096
UGRADFTE	5,304	3,071	5,942	3.335[b]	.000

[a]Denotes Z value for test of differences between proportions.
[b]Donotes t value for test of differences between means.

There were also statistically significant differences in expenditures on academic support services, the percent of faculty who were full-time and the size of student enrollment between the two types of institutions. While HWCUs have significantly higher per student expenditures on academic support services, HBCUs have a significantly higher percentage of full-time faculty than HWCU. Finally, HWCUs, on average, have nearly twice the student enrollment of HBCUs.

The results of the fitted regression equation are shown in Table 2. Overall, the explanatory variables account for 52 percent of the variation in 6-year graduation rates. In general, the variance inflation factors (VIFs) are low and suggest that multicollinearity is not a problem. The only exceptions are the relatively large VIF values for the variables UGRADFTE and $UGRADFTE^2$. However, this is as expected since $UGRADFTE^2$ is simply a transformation of UGRADFTE. Nevertheless, the coefficient for the variable UGRADFTE is highly significant and the coefficient of the variable $UGRADFTE^2$ is statistically significant at about the five percent level.

Table 2: Results of Fitted Regression Equation

Variable	b	t	Significance Level	VIF
Constant	-56.841	-2.123	.036	
HSGPA	19.742	3.970	.000	2.299
SAT	.013	.517	.606	4.405
PCTPELL	-.468	-3.428	.001	3.875
PCTFTFAC	.224	2.475	.015	1.407
ACADSUPT	.002	3.371	.001	1.130
UGRADFTE	.002	2.589	.011	12.275
$UGRADFTE^2$	-9.980E8	-1.969	.051	12.051
HBCU	23.184	5.350	.000	3.091

R^2=.518, SEE=11.0847%, n=117

The student characteristic variables HSGPA and PCTPELL are statistically significant and have the expected signs. A one fourth letter grade increase in the average high school grade point average is associated with about a five-percentage point increase in the 6-year graduation rate. A ten-percentage point increase in the number of students receiving PELL grants is associated with about a 5-percentage point decrease in the 6-year graduation rate. The median SAT or equivalent ACT score was not statistically significant. This suggests that these tests scores are not a good predictor of 6-year graduation rates at the sample institutions. This finding is consistent with other research findings. Aguinis, Culpepper and Pierce (2016) found that SAT scores were not predictive of freshman GPAs at a significant number of colleges and universities.

All three of the institutional variables were statistically significant and had the expected signs. A 10-percentage point increase in the percent of full-time faculty is associated with a 2.24 percentage point increase in the 6-year graduation rate. A

$1,000 increase in expenditure on academic and student support services per student is associated with a 2-percentage point increase in the 6-year graduation rate. Both of the terms for the size of the institution as measured by UGRADFTE had the expected signs. The term UGRADFTE is highly significant and the term UGRADFTE2 is statistically significant at the .051 level. The findings with respect to this size variable suggest that 6-year graduation rates increase as the size of the institution increases up to a point. However, after an institution reaches a certain size, 6-year graduation rates begin to decline.

Finally, the variable HBCU is highly significant. After controlling for other variables, 6-year graduation rates of African American students at public HBCUs are 23-percentage points higher than those at their peer HWCUs.

Conclusions

The findings of this study indicate that African American students at HBCUs have significantly lower high school GPA's and SAT scores and come from lower socioeconomic backgrounds than their peers at HWCUs. Despite these disadvantaging factors, the 6-year graduation rates for African American students are not significantly different between the two types of institutions. And, after controlling for differences in student characteristics and institutional inputs, HBCUs generate 6-year graduation rates that are 23 percentage points higher than those of their peer HWCUs.

These findings suggest that HBCUs are more adept at educating less well academically prepared African American students who come from lower socioeconomic backgrounds than their peer HWCUs. These findings would tend to rebut arguments by some critics of HBCUs who suggest that HBCUs no longer play a productive and meaningful role in educating African American students.

The study does have a number of limitations. First, it assumes that an institution's 6-year graduation rate is an appropriate metric for evaluating an institution's performance. In addition, it assumes that it is valid to compare the 6-year graduation rates between the two types of institutions. Finally, the study does not identify the specific factors that account for the differences in the performance of the two types of institutions.

The limitations of the study suggest a need for further research in two important areas. First, there is a need to take a more detailed look at the two types of institutions and attempt to identify what it is about HBCUs that gives them a comparative advantage in educating (as measured by generating 6-year graduation rates) less well academically prepared African American students from lower socioeconomic backgrounds. This will require a much more detailed evaluation of factors that are not readily available in public databases. In addition, there is a need to identify additional metrics with which to compare the effectiveness of the two types of institutions in educating African American students and to see how they compare using these additional measures of performance.

References

Aguinis, H., Culpepper, S., & Pierce, C. (2016). Differential prediction generalization in college admissions testing. *Journal of Educational Psychology*, 108(7), 1045-1059.

Allen, W. (1987). Black Colleges vs. White Colleges: The Fork in the Road for Black Students. *Change,* 19(3), 28-34.

Allen, W. (1992). The Color of Success: African-American College Student Outcomes at Predominantly White and Historically Black Universities. *Harvard Educational Review*, 62(1), 26-44.

Anderson, M. (2015*). A look at historically black colleges and universities as Howard turns 150.* Retrieved from the Pew Research Center website: http://www.pewresearch.org/fact-tank/2017/02/28/a-look-at-historically-black-colleges-and-universities-as-howard-turns-150/.

Anstine, J. (2013). Graduation rates at U.S. colleges and universities: A large data set analysis. *Business Education & Accreditation,* 5(2), 55-64.

Archibald, R. & Feldman, D. (2008). Graduation rates and accountability: regressions versus production frontiers. *Research in Higher Education*, 49, 80-100.

Arum, R. & Roksa, J. (2011). *Academically Adrift*. Chicago, IL: University of Chicago Press.

Campbell, K. & Fuqua, D. (2008). Factors predictive of student completion in a collegiate honors program. *Journal of College Student Retention***,** 10(2), 29-153.

Cohn, E., Rhine, S., & Santos, M. (1989). Institutions of higher education as multi-product firms: Economies of scale and scope. *The Review of Economics and Statistics*, 71(2), 284–290.

Egerton, J. (1975). Adams V. Richardson: Can separate be equal? *Change,* 6(10), 29-36.

Fryer, R., & Greenstone, M. (2007, April). *The causes and consequences of attending historically black colleges and universities*. National Bureau of Economic Working Paper No. R13036. Cambridge, MA.

Gasman, M., Baez, B., Drezner, N., Sedgwick, K., Tudico, C., & Schmid, J. (2007). Historically Black Colleges and Universities: Recent Trends. *Academe*, 93(1), 69–78.

Gasman, M., & Bowman, N., III. (2011). How to paint a better portrait HBCUs. *Academe*, 97(3), 24–27.

Ginder, S., Kelly-Reid, J. & Mann, F. (2016). *Graduation Rates for Selected Cohorts, 2007-12, Fall 2015: First Look* (NCES 2017-084). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved from: http://nces.ed.gov/pubsearch.

Kim, M. , (2002). Historically Black vs. White Institutions: Academic development among Black students. *The Review of Higher Education*, 25(4), 385-407.

Kim, M., & Conrad, C. (2006). The Impact of Historically Black Colleges and Universities on the academic success of African-American students. *Research in Higher Education.* 47(4), 399-425.

Mercer, C. (2008). *Title III and Title V of the Higher Education Act: Background and Reauthorization Issues*. Retrieved from the Congressional Research Service website: http://congressionalresearch.com/RL31647/document.php?study=Title+III+and+Title+V+of+the+Higher+Education+Act+Background+and+Reauthorization+Issues

Millea, M., Wills, R., Elder, A., & Molina, D. (2018). What matters in college student success? Determinants of college retention and graduation rates. *Education*, 138(4), 309–322.

Montgomery, R., & Montgomery, B. (2012). Graduation rates at Historically Black Colleges and Universities: An underperforming performance measure for determining institutional funding policies. Journal of Continuing Higher Education, 60(2), 93-109.

Morrison, N. (2015, July 9) Higher Education is Failing Students and Employers. *Forbes.* Retrieved from https://owl.purdue.edu/owl/research_and_citation/apa_style/apa_formatting_and_style_guide/reference_list_electronic_sources.html.

Riley, J. (2010, September 28). Black colleges need a new mission. Wall Street Journal, Retrieved from https://www.wsj.com/articles/SB10001424052748704654004575517822124077834.

Singletary, M. (2018, October 5). U.S. student loan debt reaches a staggering $1.53 trillion. *The Washington Post*, Retrieved from https://www.washingtonpost.com/business/2018/10/04/us-student-loan-debt-reaches-staggering- trillion/?noredirect=on&utm_term=.c4f81a7393a5.

Slaughter, A. (2017, May 17). The Broken Promise of Higher Education. *The Atlantic.* Retrieved from https://www.theatlantic.com/education/archive/2017/05/the-broken-promise-of-higher-education/526995/

Smith, C. (2018, January). Higher Education Is Drowning in BS. *The Chronicle of Higher Education.* Retrieved from https://www.chronicle.com/article/Higher-Education-Is-Drowning/242195.

Stewart, P. (2011). HBCU merger proposals persist despite fervent opposition. Issues in Higher Education, 27, 10-11.

Tinto, V. (1993). *Leaving college: Rethinking causes and cures of student attrition* (2nd ed.). Chicago, IL: University of Chicago Press.

U.S. Department of Education. (2018). *Integrated Postsecondary Education Data System. Retrieved from* https://nces.ed.gov/ipeds/use-the-data.

Webber, D., & Ehrenberg, R. (2010). Do expenditures other than instructional expenditures affect graduation and persistence rates in American higher education? *Economics of Education Review* 29, 947–958.

Wenglinsky, H. (1996). The educational justification of Historically Black Colleges and Universities: A policy response to the U. S. Supreme Court. *Educational Evaluation and Policy Analysis,* 18(1), 91-103.

Winkle-Wagner, R. & McCoy, D. (2018). Feeling like an "Alien" or "Family"? Comparing students and faculty experiences of diversity in STEM disciplines at a PWI and an HBCU. *Race, Ethnicity and Education,* 21(5), 593-606.

Zarifa, D., Kim, J., Seward, B., & Walters, D. (2018). What's taking you so long? Examining the effects of social class on completing a bachelor's degree in four years. *Sociology of Education.* 91(4), 290-322.

Zhang, L. & Worthington, A. (2018). Explaining Estimated Economies of Scale and Scope in Higher Education: A Meta-Regression Analysis. *Research in Higher Education,* 59:156–173.

Minecraft Economics: A Study of Wealth Inequality in a Virtual World

Calvin Blackwell, College of Charleston
Jack Carroll, Clemson University

Abstract

We examine the hypothesis that capitalist economies generate inequality by looking at the distribution of wealth in a virtual world. Employing the sandbox video game Minecraft as an example of a "pure" market economy, we chose several Minecraft economy servers and collected their respective wealth data. We found these unregulated, virtual economies to have greater wealth inequality than any single nation's "real world" economy. These results suggest that market economies generate significant "natural" inequality, and that government policies are effective in moderating the inequality-inducing impact of capitalism.

Introduction

A recurring question in economics concerns how much inequality capitalist economies create "naturally" (i.e., without government intervention). Some authors, (e.g., Kuznets, 1955) have argued that capitalist economies tend to reduce inequality, while others (Ricardo, 1911; Marx 1887; and Piketty, 2014) have argued the opposite. The issue is a difficult one to address empirically for at least two reasons. First, the unit of analysis (typically a country) is subject to many factors that can affect inequality. For example, numerous parts of the tax code in modern economies exist to reduce inequality; however, given the complexity of modern tax codes, it is challenging to determine how individual requirements impact inequality. Relatively sparse data means there is a high ratio of explanatory factors to data points, making empirical analysis challenging. Second, although there is relatively good data on income, in many countries it is difficult to gather data on wealth. For example, there is no complete data set on the wealth distribution in the United States; instead, researchers must rely on incomplete data sets that sometimes generate conflicting conclusions (Kopczuk, 2015).

This paper attempts to address how inequality evolves in market economies by addressing the two common challenges listed above. We examine the distribution of wealth in a market environment with minimal government intrusion, and in an environment in which we can know exactly how much wealth each individual holds. Although these virtual economies come from video games, the richness of the economic interactions that are available to players, combined with the low regulatory environment and extremely high-quality data, allow us to begin to draw some conclusions regarding the evolution of wealth distribution. Our results indicate that these "pure" market economies have higher wealth inequality than low inequality economies (like Japan), high inequality economies (like the United States), and the current global measure of inequality. We interpret this finding as support for the claim that markets do tend to generate inequality, but also as evidence that governments' policies to reduce inequality are effective.

Methodology

To assess the "natural" evolution of inequality under "pure" capitalism, one needs to find an economy with no redistributive policies. Such economies do not exist in the real world, but sometimes exist in virtual worlds, like Second Life or World of Warcraft. The virtual world that will be used in this study is Minecraft - a sandbox video game created in 2011 by the Swedish programmer Markus "Notch" Persson and published by Mojang (now available at minecraft.net).

Readers may be skeptical that data from a virtual world is useful for understanding "real" economies. However, Castronova et al., 2009 and Castronova, Knowles, and Ross, 2015 argue that virtual economies tend to behave like other, "real" economies. In fact, researchers have looked at virtual economies to examine a number of economic phenomena, including GDP identities and the quantity theory of money (Castronova et al., 2009) and the evolution of unregulated capital markets (Bloomfield and Cho, 2011). Data from these virtual economies is consistent with standard economic theory.

The Minecraft video game allows players to meet virtually on "servers," where a wide variety of activities take place. Servers are virtual locations, hosted by one or more individuals. These hosts are able to control a variety of parameters regarding how Minecraft is played on their server. Players can chat or play games with one another on the servers. When players start Minecraft, they can choose among servers based on several characteristics. The particular type of server that we are interested in is labeled "economy" by Mojang, and has the following basic set of rules. The primary goal of the players on the server is to advance in rank, which can be purchased with gold, which functions as these economies' unit of account and exchange.

Players interact primarily in two types of environments: mines and markets. In mines, players can employ tools to "mine" or collect resources like diamonds or iron ore. These resources have a degree of fungibility; some can be turned into capital goods like tools or armor, and all resources can be bought and sold in markets. In some types of mines, players merely collect resources. In other mines, labeled PVP ("Player vs Player"), players are able to interact with each other violently. For example, in a PVP mine player A could attack player B and take all the resources on B's person. In markets, players can buy and sell resources and are able to use gold to purchase higher ranks. Unlike in mines, in markets the rules of the game do not allow for violence or involuntary exchange; that is, property rights are perfectly enforced. (Here we see the application of Lessig's observation that in virtual economies, "code is law" (Lessig, 2008).) Although each server may have custom rules, the basic observation is that each server has an economy, with a standardized unit of exchange, and with well-defined rules for making market transactions. Importantly, there are no redistributive rules - each player is responsible for his/her own resources, and transfers of resources from one player to another are never made by individuals running the server.

We wish to examine the distribution of wealth. The natural analog of wealth in Minecraft is gold – as mentioned previously, it is the unit of account and exchange in the game, and the incentives in the game are for players to maximize the acquisition of gold so as to advance in rank. We argue that measuring wealth in gold in Minecraft is similar to measuring wealth as the value of assets in dollars in the real world.

Data Collection

Each player has a continuously updated inventory of goods, armor and gold. This information is stored on the Minecraft server, and so can be accessed by both the individual player and the host. Some servers allow players to view the gold balances of other players with in-game commands. Our data was collected using the in-game command "/baltop", which shows the current gold balance of every player in that server. The players were ranked based on their numerical wealth value, i.e. larger values first, which were displayed in on pages, with nine wealth values per page. The quantity of pages depends on the size of the server. The nine servers were chosen randomly from http://minecraft-mp.com/; however, the "economy tag" had to be the foremost tag detailing the server. The data (on 17,676 players) was collected between July 12, 2016 and September 15, 2016.

Gini Coefficient

We use the Gini coefficient to measure wealth inequality. Its value ranges from 0 to 1, 0 the value of perfect equality and 1 the value of maximum inequality. Numerous formulas for the Gini coefficient abound, but given the nature of the data we wish to analyze, we will use the following formula (from (Deaton, 1997)):

$$Gini = \frac{N+1}{N-1} - \frac{2}{N(N-1)\mu}\left(\sum_{i=1}^{N} P_i Y_i\right) \qquad (1)$$

where N is the number of observations, μ is the mean wealth, P_i is the income rank P of person i, with an income of Y, in a way that the poorest individual receives a rank of N and the richest a rank of 1.

Results

Table 1 gives the descriptive statistics for all of the servers. The number of players across servers varies considerably, as does average wealth.

Table 1: Descriptive Statistics

Server	Charlie	Evisea	Ezure	Gcraft	Norse	Odin	Pcraft	Sky Srvrs	Tempus
# players	387	1333	1719	349	1223	3422	759	2770	5714
AVG wealth	2.5E1	5.1E3	1.5E6	8.1E4	6.5E4	8.8E6	2.0E9	1.6E4	9.0E6
STD DEV wealth	5.5E1	4.1E4	8.8E6	3.1E5	8.3E5	3.2E8	3.5E10	3.3E5	6.7E8
Max wealth	4.5E2	8.5E5	3.8E7	3.0E6	2.0E7	1.3E10	6.2E11	1.4E7	5.0E10
Gini	0.9262	0.9172	0.9794	0.8791	0.986	0.9995	0.9987	0.9503	0.9998

This second observation comes from the fact that the "price level" of the server is set idiosyncratically by the server's owner - some servers have high price levels, (reflecting a large money supply), while other servers have lower price levels (again, reflecting a lower money supply). As is the case in the real world, the distribution of wealth is highly positively skewed, with the standard deviation of wealth typically being an order of magnitude greater than the mean. We see that the levels of inequality are high, with only one server having a Gini coefficient less than 0.9.

In Table 2, we provide Gini coefficients from Davies et al., 2009 on comparable economies, including their estimate of world-wide inequality. We also present data on inequality generated in the agent-based simulation Sugarscape (Epstein and Axtell 1996). Note that the Minecraft economies are more unequal than even the least equally distributed national economy (Namibia), and are more unequal than the world economy. The Minecraft economies are substantially more unequal than developed economies with low inequality (e.g., Japan) and high inequality (e.g., the United States), which have significant social safety nets and policies to mitigate inequality.

Table 2: Sample Gini Coefficients.

Unit	*Gini*
Minecraft Average	0.96
World	0.89
Namibia	0.85
United States	0.8
Sugarscape (inheritance)	0.75
Japan	0.54
Sugarscape (no inheritance)	0.5

Discussion

We have presented evidence that virtual economies lacking redistributive policies appear to generate more inequality than real economies. This result is suggestive of the conclusion that unfettered capitalism has the potential to generate substantial inequality. However, our findings must be tempered due to a number of potentially confounding factors. First, one must ask how well does a Minecraft economy mimic a real economy? Although the Minecraft marketplace allows for rich interactions between players, it is incredibly simple when compared to the real world. For example, there are no capital or insurance markets, and property rights are cartoonish, either fully enforced or not at all. In addition, the level of violence between players is much higher than in most real economies. Also, the ability of players to form coalitions or other sorts of contracts is absent. These factors may contribute to the high levels of inequality observed.

Second, one must consider that the level of wealth inequality may have been (advertently or inadvertently) designed into the game. Games in which the outcome is the same for all players may not be considered entertaining, leading game designers to make decisions that increase inequality.

We can gain some more insight by considering how the Minecraft economies compare to the Sugarscape economy (Epstein and Axtell 1996). In their simple agent-based simulation, agents collect sugar, which is necessary for the agent to live and to reproduce. The agents can trade sugar, and the amounts of sugar held by agents can be observed. The Sugarscape and Minecraft economies are very similar, with the major difference being that in Minecraft economies, resources can be taken through violence. In their simulation, Epstein and Axtell generated only moderate amounts of inequality in sugar; however, they also showed how sensitive this inequality was to time. That is, in the Sugarscape economy initially all agents had similar levels of wealth, and as the agents were allowed to harvest and trade, more inequality arose. It seems likely that a similar process arises on the Minecraft servers. If we use the number of players on the server as a proxy for the length of time the server has been in existence, we see that larger numbers of participants are positively correlated ($r=0.55$) with the Gini coefficient.

In one Sugarscape simulation, the authors allowed for agents to inherit wealth from their predecessors. This change increased inequality substantially, increasing the Gini coefficient from 0.5 to 0.75. Minecraft servers, which do not allow for inheritance, but do allow for violence, exhibit bigger Gini coefficients than Sugarscape with inheritance, suggesting that violence and/or poorly defined property rights may have a larger impact on inequality than inheritance.

Acknowledgment

References

Bloomfield, Robert J., and Young Jun Cho. 2011. "Unregulated Stock Markets in Second Life." *Southern Economic Journal* 78 (1): 6–29.

Castronova, Edward, Isaac Knowles, and Travis L. Ross. 2015. "Policy Questions Raised by Virtual Economies." *Telecommunications Policy*, SPECIAL ISSUE ON THE GOVERNANCE OF SOCIAL MEDIA, 39 (9): 787–95.

Castronova, Edward, Dmitri Williams, Cuihua Shen, Rabindra Ratan, Li Xiong, Yun Huang, and Brian Keegan. 2009. "As Real as Real? Macroeconomic Behavior in a Large-Scale Virtual World." *New Media & Society* 11 (5): 685–707.

Davies, James B., Susanna Sandstrom, Anthony Shorrocks, and Edward Wolff. 2009. "The Global Pattern of Household Wealth." *Journal of International Development* 21: 1111–24.

Deaton, Angus. 1997. *The Analysis of Household Surveys*. Baltimore, Maryland: Johns Hopkins University Press.

Epstein, Joshua M., and Robert Axtell. 1996. *Growing Artificial Societies: Social Science from the Bottom Up*. Complex Adaptive Systems series. Washington, D.C Brookings Institution; Cambridge and London: MIT Press.

Kopczuk, Wojciech. 2015. "What Do We Know about the Evolution of Top Wealth Shares in the United States?" *Journal of Economic Perspectives* 29 (1): 47–66. https://doi.org/10.1257/jep.29.1.47.

Kuznets, Simon. 1955. "Economic Growth and Income Inequality." *The American Economic Review* 45 (1): 1–28.

Lessig, Lawrence. 2008. *Code*. Version 2.0. New York: Basic Books.

Marx, Karl. 1887. "Capital: Volume One." 1887. https://www.marxists.org/archive/marx/works/1867-c1/.

Piketty, Thomas. 2014. *Capital in the Twenty-First Century*. 3rd edition. Belknap Press.

Ricardo, David. 1911. *On the Principles of Political Economy and Taxation*. 2004, reprint, London: Dover Publications.

The Liquidity and Performance of SRI Funds

Jennifer Brodmann, California State University Dominguez Hills
Phuvadon Wuthisatian, Hastings College

Abstract

Socially Responsible Investment (SRI) is steadily gaining popularity throughout the world as an alternative means of investment that includes ratings based on environmental, social, and governance (ESG). Previous studies provide mixed results of whether SRI funds outperform conventional funds. We examine SRI asset performance in comparison to conventional assets using the KLD 400 Index over a time span of over 25 years from 1990 until 2015. We use the S&P market index to predict market excess returns using idiosyncratic volatility and determine whether SRI investments have lower liquidity than their conventional counterparts. We examine the required return that SRI investors expect in order to maintain their holdings in SRI stock. Our findings reflect a significant difference in liquidity and turnover ratio between SRI and non-SRI firms. This suggests that there is a stigma against investing in SRI, with investors preferring to invest in non-SRI firms.

Introduction

Hayes (2005) concludes that "Socially responsible investment represents one of the fastest growing equity investments." This new type of investment is gaining traction and several market indices include stocks that adhere to socially responsible business practices and are thus considered socially responsible investments (SRI). Schueth (2003) states that there are two main motivations for SRI investors, which is the need to invest in assets that align with their 'personal values and priorities' and the other is the need to invest in assets that 'support and encourage improvements in quality of life.'

Typically, investors that are more focused on returns and diversification may be wary of SRI investment from this additional SRI criteria screening, which may make SRI assets less liquid compared to conventional assets. Previous studies contend that funds that screen for ethical and moral characteristics may impose additional constraints to investors that are interested in wealth maximization (Grossman and Sharpe, 1986; Hall, 1986; Rudd, 1981). Yet, the results in Hill, Ainscough, Shank, and Manullang (2007) reflect that firms that engage in socially responsible business practices may increase their firm value in the long-term. This finding is also supported in Shank, Manullang, and Hill (2005), who find that the market incorporates SRI characteristics into prices in the long-run. We examine this phenomenon further to determine whether SRI firms outperform non-SRI firms, if either SRI firms are as liquid as non-SRI firms, and if SRI firm funds can be used for diversification during financial crises. This contributes to the literature by exploring whether there is investor bias against SRI firms, because the firms incorporate environmental, social, and governance factors, instead of just focusing on financial indicators into their decision-making.

Friede, Busch, and Bassen (2015) conduct an extensive study and examine the evidence stated in over 2000 empirical studies on environmental, social and governance criteria (ESG) and corporate financial performance (CFP). They state that the majority of the studies find a positive relationship between ESG and CFP. In contrast, other studies have found weak evidence on whether SRI funds out- or underperform conventional funds. Yet, studies find that SRI funds have lower flow volatility compared to conventional funds (Bollen, 2007). There is no clear evidence whether conventional funds outperform SRI funds. This motivates our interest to test for the performance of conventional funds compared to SRI funds. Using a comprehensive dataset spanning from 1990 to 2015, we argue that, in fact, the SRI and non-SRI firms actually perform in similar manners. Then, it must be other reasons behind why investors are investing in conventional funds instead of SRI funds. Lean, Ang, and Smyth (2015) also provide an empirical study of SRI funds in Europe and North America and conclude that the SRI funds perform better than market benchmark.

Bollen (2007) provides empirical results indicating that SRI assets yield lower volatility than non-SRI assets. Despite lower volatility, the evidence in some literature shows that SRI investments may have lower liquidity due to investor bias or lack of SRI fund information for investors (Van Liedekerke, De Moor, and Vanwalleghem, 2007). There may be investor stigma against investing in SRI funds because SRI firms factor in ecological, social, and governance factors into their investment decision making instead of just focusing on maximizing stock value through financial criteria. This can be seen in the case of Asia Pacific markets, where addition and deletion of stocks in sustainability indexes can result in a decrease in returns, increases in trading volume and idiosyncratic risk, which may indicate that investors have a negative reaction towards stocks that are categorized as sustainable (Roca, 2013).

The impact of firms implementing corporate social responsibility practices has been widely studied in the previous literature. Some findings justify the stigma of CSR being unable to achieve positive outcomes. Aupperle, Carroll, and Hatfield (1985) conclude that there is no relationship between corporate social responsibility and profitability. In contrast, Cochran and

Wood (1984) find a positive correlation between CSR and financial performance. Barnea and Rubin (2010) study managements' practices in investing in CSR and find that overinvestment in corporate social responsibility activities may happen when insiders are focused on building their reputations as good global citizens, with having the firms bear the cost of the investment.

Previous studies intimate but do not provide substantive evidence that SRI investors are open to accepting underperformance to adhere to social or ethical values (Renneboog, Ter Horst and Zhang, 2008). Some studies examine the performance, risk-return, and volatility of SRI funds, yet there are very limited studies on SRI fund liquidity. Fernando, Sharfman, and Uysal (2009) examine the liquidity of SRI investing in green businesses and their findings suggest that greener firms have higher stock market liquidity. This may hold true for businesses that adhere to environmental SRI standards, but will this hold true for firms that also adhere to social and governance standards? We expand on this topic by examining the liquidity of SRI stocks from the KLD 400 index, which encompasses investments of all ESG factors. We hypothesize that if SRI stocks yield higher liquidity, then we prove that, in fact, SRI investments can be used for investors not only to receive a substantial amount of returns but also can provide attractive investments for them. We construct the liquidity measure using Amihud (2002) illiquidity measure as well as the turnover rate (Amihud and Mendelson, 1986) to test for liquidity between these funds. The result shows that, consistent with most literature, SRI firms actually are less liquid than non-SRI firms. Furthermore, we also confirm the result using turnover rate as the proxy for liquidity measure that SRI funds are less liquid than conventional funds. The results provide an intuition that SRI assets as alternative investments may not provide liquidity in trading strategy, but rather investors use SRI assets to diversify their investment decisions.

There is an argument that if SRI funds provide diversification benefits, therefore we must expect to see the returns of SRI funds outperform those of conventional funds. Previous studies have found that firms that practice CSR may be able to fare better during times of crisis (Aravelo and Araving, 2010; Schneitz and Epstein, 2005). We examine this further by measuring the performance of SRI firms and non-SRI firms to determine whether SRI firms were able to weather the Global Financial Crisis better than non-SRI firms. Our results show that SRI firms provide higher returns during the financial crisis period compared to non-SRI firms. This evidence indicates that investors invest in SRI assets to diversify their portfolios during the financial crisis.

How socially responsible firms weather crises is a topic of interest as well. Schneitz and Epstein (2005) find that socially responsible firms react negatively to crises, such as the failure of the Seattle World Trade Organization (WTO) ministerial meeting. SRI Firms that may be able to weather crises better than firms than non-SRI firms. Arevalo and Aravind (2010) find that firms that engage in CSR efforts well less affected by the GFC than firms that are not engaged in CSR efforts using the United National Global Compact. Giannarakis and Theotokas (2011) determine that CSR performance increase pre- and post-GFC. We examine this in the context of whether SRI firms fared better from the GFC using the KLD 500 index. Souto (2009) provides recommendations of executing CSR models from a threat into an opportunity during crisis periods.

Ang et al. (2006) and Fu (2009) test for the stock return in the U.S. market by observing the idiosyncratic risk and sort portfolios based on the idiosyncratic risk. Idiosyncratic risk, as defined by the literature, is the error term of the regression, which helps explain the change in the stock movement in which it is not correlated with the market risk. Their results show that using portfolio approach sorting can yield significant gains higher than the market benchmark. We estimate using the GJR-GARCH model to capture the leverage effect for both SRI and non-SRI firms. Engle and Ng (1993) test with various ARCH GARCH types and conclude that GJR-GARCH (Glosten, Jagannathan and Runkle, 1993) is incorporated with the new information of the U.S. stock price.

We contribute to the current literature in the following ways: (i) by estimating the volatility of SRI versus conventional stocks (ii) by finding the excess return of SRI and conventional stocks as well as the market return, and (iii) by comparing the liquidity of SRI stocks to conventional stocks measuring liquidity through the illiquidity and stock turnover ratio measures in Amihud (2002); (iv) sorting portfolios into quantiles to find the patterns from the conventional stocks and (v) by testing how SRI and non-SRI stocks reacted to the Global Financial Crisis (GFC); and (vi) determining the required return that SRI investors expect in order to maintain their holdings in SRI stock.

Data and Methodology

SRI and non-SRI Data

We collect daily stock price data for both the conventional and SRI stocks from the Center for Research in Security Prices (CRSP) using sharecode 10 and 11 to identify stocks that trade on the New York Stock Exchange (NYSE), American Market Exchange (Amex), and Nasdaq. We exclude penny stocks and stocks that have fewer than five years of data. Also, we winsorize 0.5% for each tail to avoid potential outliers (See. Hoberg and Phillips, 2010). We use KLD data as well as the MSCI KLD 400 Social Index to classify which firms and their subsequent stocks are considered SRI stocks. The sample period is 25 years

from 1990 until 2015 to capture a sample that is before and after the Global Financial Crisis. The sample consists of 1,093,957 observations with a total number of 16,933 firms, with 374 SRI firms and 16,619 non-SRI firms.

We use value weighted to measure the excess return for these funds. Also, the market return is value weighted return (See Fama and French, 2016; Banz, 1981; Pastor and Stambaugh, 2003). We find that there the SRI and non-SRI firms are similar in terms of performance, with SRI firms having 3.03 percent excess return and non-SRI firms having a 3.67 percent excess return. In addition, the volatility is also quite similar, with SRI firms having a standard deviation of 2.02 percent and Non-SRI firms having a standard deviation of 2.13 percent. Both funds, however, outperform the market index.

The sectors are classified by KLD 400 database. Real Estate provides the highest return of 5.14% while Consumer Staples has the lowest mean return of 0.33%. However, when looking at the size of the standard deviation of each sector, we find that each sector provides similar size of the volatility. Then, the plausible explanation for investors to use SRI as an alternative investment is to diversify their portfolios since SRI funds provide a stable amount of volatility for all sectors. We also report the market value of SRI funds by sectors. Information Technology has the highest market value of asset reflecting that the size of firms in this sector is bigger than any other since it requires a substantial amount of investment to get into the Information Technology industry.

Liquidity Measure

We measure liquidity as in Amihud (2002), which states that "illiquidity if the relationship between the price change and the associated order flow or trading volume." Stock illiquidity is considered as the average ratio of the absolute daily return and the trading volume in dollars that day.

$$\text{stock illiquidity} = \frac{|R_{i,yt}|}{VOLD_{i,yt}} \tag{1}$$

where $R_{i,yt}$ is the return on stock i on day t of year y, $VOLD_{i,yt}$ is the respective daily trading volume in terms of dollars. Equation 1 reflects the absolute percentage change in price for each dollar of the daily trading volume or the impact of the daily price of the order flow. To determine stock illiquidity for each year y, the measure for each stock i is determined as the average.

$$ILLIQ_{i,y} = \left(\frac{1}{D_{i,y}}\right) \sum_{t=1}^{D_{i,y}} \frac{|R_{i,yt}|}{VOLD_{i,yt}} \tag{2}$$

where $D_{i,yt}$ is the number of days where there is available data for stock i in year y. Stock liquidity is also measured through the turnover ratio (see Amihud and Mendelson, 1986; Amihud, 2002; Amihud, Mendelson, and Pedersen, 2006; Datar, Naik, and Radcliffe, 1998), stated as follows:

$$TURN_{i,y} = \left(\frac{1}{D_{i,y}}\right) \sum_{t=1}^{D_{i,y}} \frac{VOLSHS_{i,yt}}{NSHRS_{i,yt}} \tag{3}$$

where $VOLSHS_{i,yt}$ is the trading volume in terms of the shares of stock i on day t in year y, $NSHRS_{i,yt}$ is the number of outstanding shares of stock i on that day. Amihud and Mendelson (1986) find that asset illiquidity costs are negatively related to trading frequency, in equilibrium. This occurs from liquid stocks being allotted to investors that execute less frequent trades who write off the illiquidity cost across a longer time period, which reduces the illiquidity cost of the asset. Therefore, we expect that stocks with a lower turnover ratio are less liquid and that stocks that have a higher turnover ratio are more liquid.

We report the liquidity for both SRI and non-SRI firms. As expected, using Amihud's illiquidity the SRI firms are less liquid than non-SRI firms. Higher illiquidity means the lower liquidity while the higher turnover ratio indicates the higher liquidity (See Brunnermeier and Pedersen, 2008). Although, SRI assets can be used to diversify portfolios, due to the constraints imposed to meet social standards making them less liquid than conventional funds.

Regression Analysis

We conduct the regression test to see whether the liquidity measure can influence the performance of both SRI and non-SRI funds. The regression is determined by the following model:

$$er_{i,t} = \alpha_{i,t} + \beta_1 liq_{i,t} + \beta_2 turn_{i,t} + \beta_3 X_{i,t} + \varepsilon_{i,t} \tag{4}$$

where $er_{i,t}$ is the excess return of stock i and time t, $liq_{i,t}$ is Amihud's illiquidity measure for stock i at time t, $turn_{i,t}$ is the turnover ratio for stock i at time t, and $X_{i,t}$ is the vector of economic variables to control for the change in excess return of stock i at time t. The economic variables are VIX, T-bill, CPI, and bond yield. These variables are supported in the literature that they influence the change in excess return of the U.S. equity. We also include the log of firm size to control for the regression model (See Adrian and Shin, 2010; Gorton and Metrick; 2012, Brunnermeier and Pedersen, 2003). We multiply the Amihud's illiquidity measure with -1 to get the liquidity measure. The intuition of including these two liquidity measures is that these variables must be able to capture the change in the excess return for stock in both SRI and non-SRI funds.

We also test further for the impact of financial crisis on the firm performance. The dummy variable equal to 1 indicating a financial crisis, and zero otherwise. The time period is March 2008 to January 2010 as VIX spread shows the sign of high volatility. The VIX spread, defined by the Chicago Board of Options Exchange (CBOE), indicates that the implied volatility of the market index. The higher the spread means that the higher the expected volatility and investor's fear. The model is estimated as the following:

$$er_{i,t} = \alpha_{i,t} + \beta_1 illiq_{i,t} * \left(Dummy_t^{crisis}\right) + \beta_2 turn_{i,t} * \left(Dummy_t^{crisis}\right) + \beta_3 X_{i,t} * \left(Dummy_t^{crisis}\right) + \varepsilon_{i,t} \qquad (5)$$

Portfolio Construction - Idiosyncratic Volatility

We now turn our analysis to test for the portfolio performance between SRI and non-SRI firms. Ang et al. (2009), and Fu (2009) argue that the idiosyncratic risk helps to explain the change in the stock movement, which is uncorrelated with the market risk. They also suggest that idiosyncratic risk can be used to predict the short-term performance of stocks. Motivated by this finding, we are interested in testing whether sorting portfolios based the idiosyncratic risk can be able to explain the difference in stock returns of SRI and non-SRI firms.

We use a 5-factor model to predict the stock return as follows:

$$er_t = \alpha_i + \beta_i MKT_t + s_i SMB_t + h_i HML_t + c_i CMA_i + r_i RMW_i + \varepsilon_{i,t}, \varepsilon_{i,t} \sim N(0, \sigma_{it}^2) \qquad (6)$$

Where er_t is the daily excess U.S. dollar return of stock i, MKT_t is the value-weighted of local market portfolio over the one-month T-bill rate, SMB_t is the return of the smallest one-third of local firm minus the return of the largest one-third of local firm characterized by the market capitalization, and HML_t is the return of the highest one-third of book-to-market ratio minus the return of the lowest one-third of the lowest book-to-market ratio, CMA_i (conservative minus aggressive) is an investment factor, and RMW_i (robust minus weak) is a profitability factor. The idiosyncratic volatility is measured by the standard deviation of the residual, $\varepsilon_{i,t}$, after the estimation from the regression model.

Once we determine the regression model, we now modify the idiosyncratic risk using GJR-GARCH (Glosten, Jagannathan, and Runkle, 1993). GJR-GARCH is used in order to capture the leverage effects of the stock returns, which is supported by Engle and Ng (1993) that they test all ARCH and GARCH-type models and find that GJR-GARCH is the best so far to predict the short-term predictability (see Ang et al., 2006,2009; Fu, 2009). The model specification is estimated as:

$$\sigma_{i,t}^2 = w + \sum_{i=1}^q [a_i + \gamma_i I_{(\varepsilon_{t-1<0})}] \varepsilon_{t-1}^2 + b_i \sigma_{t-1}^2 \qquad (7)$$

Equation (8) or GJR-GARCH (p,q) model, where p and q defined as the number of lags of auto-regression and moving average. The purpose of using this GJR-GARCH is to estimate the conditional variance, σ_{it}^2. The modification of the conditional variance as described in the equation is to capture the possible shocks that occur from the lagged period.

SRI Return Predictability

The argument over the return predictability has been long discussed in the literature (See. Cenedese et al., 2014; Moreira ad Muir, 2017, Fama and French, 2016). However, most of the literature focuses on conventional stocks and SRI firms are not being tested. We explore the predictability of SRI firms. We hypothesize that the SRI firms would depict some certain information to predict the movement of the return.

Fama-MacBeth (1973) suggests that two-step regression is needed to determine the coefficients of risk-loading factors. We use MKTRF, SMB, HML, CMA, and RMW as risk-loading factors since these variables are mainly used in the literature. See. Ang et al. (2006), Ang et al. (2009), Fu (2009), Fama and French (2016). Once we determine the coefficients, we can run a regression to test for the return predictability. The model is presented as the following:

$$er_{i,t+1} = \lambda_0 + \hat{\beta}_i \lambda_t + \mu_i X_t + \theta_i Z_{i,t} + \alpha_{i,t+1} \tag{8}$$

where $\hat{\beta}i$ is a vector of the coefficients estimated from the first step (MKT, SMB, HML, CMA, and RMW), and X_t is a vector of economic variables, and θ_i is the vector of control variables (Idiosyncratic factor, firm size, and liquidity factors).

The risk loading factors are collected from the Kenneth French website from http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html. Fama and French (2016) provide empirical evidence of anomalies using a five-factor model. They report that these factors can explain the average return. The independent variables of equation (8) are explained in the previous section. We include the economic factors to help to explain the potential portfolio prediction. We also add firm size, liquidity factors (Amihud's illiquidity and turnover ratio), and idiosyncratic factor to control for the SRI return prediction.

Results

Regression Results

We show the result from equation (5). As expected, the change in excess returns for both SRI and non-SRI firms are affected by the economic variables and liquidity measures as all variables are significant. The liquidity measures are, however, somehow different between these funds supported the literature that SRI funds are less liquid than non-SRI funds. The presence of lower liquidity in SRI funds making them to be used as alternative investments since investors are looking to diversify their portfolios by investing in non-tradition funds.

The VIX spread (VIX_t) provides an insightful explanation that although non-SRI funds are more liquid than SRI funds, non-SRI assets are more volatile as the VIX spread is greater than that of non-SRI. The other economic variables, namely T-bill ($T - Bill_t$), CPI (CPI_t), and Bond Yield ($Bond\ Yield_t$) for both SRI and non-SRI funds have the same signs indicating that these funds are affected by the same economic variables. This result supports the literature that these funds provide similar characteristics. Only the presence of liquidity makes investors differentiate their investment strategies.

Financial Crisis

We then test further for the presence of the financial crisis on the impact of changing in excess returns for SRI and non-SRI assets. The argument is that if SRI funds are used to diversify portfolios, then SRI funds must provide higher liquidity than non-SRI funds. During the non-financial crisis period, the result is consistent with what we find in the previous section that generally non-SRI firms are more liquid than non-SRI firms. Also, other variables have consistent coefficients as we find in the base regression model.

During the financial crisis, the VIX spread is now higher than during non-financial crisis due to the tighter funding constraints and the presence of the investment risks. Mancini et al. (2013) provide empirical evidence that during the financial crisis period the funding constraints get tighter due to the change in economic structure (see also Banti and Phylaktis (2015) and Brunnermeier and Pedersen (2009)). The VIX coefficient for SRI firms is 0.0232 compared to non-financial crisis period of 0.177 while the VIX coefficient for non-SRI firms is severely affected by the presence of financial crisis period from 0.0092 to 0.0332. The result indicates that SRI firms can withstand the financial crisis better than non-SRI firms since most SRI firms are large and mature.

Amihud's illiquidity and turnover ratio provide the consistent result with our hypothesis. SRI firms are more liquid during the financial crisis than non-SRI firms (0.130 and 0.167 for SRI, and 0.0032 and 0.0154 using LIQ and TURN respectively). Then, investors invest in SRI firms to diversify their portfolios during the financial crisis to provide them a better position in liquidity of their portfolios. Also, the result suggests that the Amihud's illiquidity, as well as turnover ratio, are able to capture the change in excess return since they are statistically significant.

Portfolio – Idiosyncratic Volatility

Ang et al. (2009) and Fu (2009) suggest the use of idiosyncratic volatility to capture the short-term movement of stocks and that portfolios sorted by idiosyncratic volatility can yield a greater return. We start our analysis for SRI firms. The portfolios sorted by idiosyncratic volatility provide the pattern that the higher volatility (portfolio 5 – 0.0701) provides a higher return (0.0373). The Shape ratio also shows a similar result with the most volatile portfolio or portfolio 5 has a Sharpe ratio of 0.5319 while portfolio 1 or the least volatile has a Sharpe ratio of 0.4943. The predictive pattern of the portfolios can be explained that

the presence of idiosyncratic risk can influence the return of the portfolios. Investors expect to receive higher returns to compensate to the higher volatility.

We then test further for non-SRI firms to see whether they provide the consistent result with SRI firms. Non-SRI firms indicate the consistent results as evidence shows that the most volatile portfolio (portfolio 5) has the highest return while the least volatile portfolio (portfolio 1) has the lowest return. Then, sorting portfolio based on idiosyncratic volatility can indicate the change in returns, in which investors trade off the higher risk to receive the higher return.

The evidence of this section, however, does not differentiate the differences between SRI and non-SRI firms. The results only indicate the pattern of risk and returns between these funds are comparable, except for portfolio 5. The portfolio 5 or the most volatile portfolio for SRI has Sharpe ratio of 0.5319 while for non-SRI has 0.6159. The distinctive difference between these funds can be explained that conventional stocks provide greater returns than SRI funds since SRI firms must meet the social standards, which are costly to operate and maintain, and severely affect the returns. Another explanation is that investors seek to maximize their investment by taking higher volatile stocks. Then, it results in that portfolio 5 for non-SRI firms provides a substantial amount of return higher than that of SRI firms.

Return Predictability

The results indicate that the presence of risk-factor loadings can help to explain the expected excess return with all these of factors being statistically significant. These factors correspond to the change in the return as suggested by Fama-French (2016) that these risk-loading factors can capture the change in short-term of stock movement. We also find that the signs for these factors are consistent with what reported by Fama-French (2016) Indicating that SRI firms behave in the same manner as in the conventional stocks.

Then, the critical part of this section is whether liquidity measures can predict the change in the return of SRI firms. We find that both LIQ and TURN are statistically significant indicating that these variables can be used to explain the change in returns of SRI firms. The result is confirmed even adding one variable at a time. However, the size of these liquidity measures is not substantially large as expected. These measures add up to approximately 16% (5.9% + 10.1%). Then, the result can be argued that although the liquidity is the important piece of the SRI funds, the predictability of these assets remains consistent with conventional assets, which can be explained by risk-loading factors. Investors prefer to have a diversified portfolio by partially investing in SRI funds while the majority of their investments is in the conventional funds. Then, this may be able to explain the small size of liquidity measures in SRI return predictability.

Then, we move on to the economic variables. We find, however, that these variables cannot be used to explain the change in SRI stock returns. Due to the differences in the characteristics to conventional stocks, the plausible explanation is that these SRI firms do provide some certain degree of creditability to investors and financial institutions, in which firms are not influenced by the other economic factors. These economic factors influence conventional stocks since they are mainly focusing on profit maximization not to fulfill social welfare.

Conclusion

This study provides an initial analysis of the liquidity of socially responsible stocks in comparison to conventional stocks. We find a significant difference in liquidity and turnover ratio, which indicates that SRI firms are less liquid than conventional firms. This may suggest that there is a stigma against investing in SRI, with investors' preference in investing in non-SRI firms. Investors that are not willing to forgo returns for socially responsible business practices may still be skeptical about investing in SRI. In Spain, investors state that more information is needed on the social, environmental, and governance criteria for them to consider investing in SRI firms and products (Escrig-Olmedo, Muñoz-Torres, and Fernández-Izquierdo, 2013).

In the future, we will explore whether and if so how much of an effect illiquidity has on the performance of SRI assets. This study examines liquidity based on the methods used in Amihud (2002), which measures liquidity through the illiquidity and turnover ratio. Other studies examine liquidity with different means of measuring liquidity. This could be conducted to provide further analysis of SRI asset liquidity. In addition, examining SRI stocks from other indexes and in other financial markets could be a means of extending the literature on this topic. Also, investigating investor demand in SRI would bring further insight as to what investors need to consider investing in SRI.

Acknowledgment

their feedback in improving this paper.

References

Adrian, T., & Shin, H. S. (2010). Liquidity and leverage. *Journal of Financial Intermediation*, *19*(3), 418-437.

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets*, *5*(1), 31-56.

Amihud, Y., & Mendelson, H. (1986). Asset pricing and the bid-ask spread. *Journal of Financial Economics*, *17*(2), 223-249.

Amihud, Y., Mendelson, H., & Pedersen, L. H. (2006). Liquidity and asset prices. *Foundations and Trends® in Finance*, *1*(4), 269-364.

Ang, A., Hodrick, R. J., Xing, Y., & Zhang, X. (2006). The cross-section of volatility and expected returns. *The Journal of Finance*, *61*(1), 259-299.

Ang, A., Hodrick, R. J., Xing, Y., & Zhang, X. (2009). High idiosyncratic volatility and low returns: International and further US evidence. *Journal of Financial Economics*, *91*(1), 1-23.

Arevalo, J. A., & Aravind, D. (2010). The impact of the crisis on corporate responsibility: the case of UN global compact participants in the USA. Corporate Governance: The international journal of business in society, 10(4),

Aupperle, K. E., Carroll, A. B., & Hatfield, J. D. (1985). An empirical examination of the relationship between corporate social responsibility and profitability. *Academy of Management Journal*, *28*(2), 446-463.

Banti, C., Phylaktis, K., & Sarno, L. (2012). Global liquidity risk in the foreign exchange market. *Journal of International Money and Finance*, *31*(2), 267-291.

Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics*, *9*(1), 3-18.

Barnea, A., & Rubin, A. (2010). Corporate social responsibility as a conflict between shareholders. *Journal of Business Ethics*, *97*(1), 71-86.

Bollen, N. (2007). Mutual fund attributes and investor behavior. *Journal of Financial and Quantitative Analysis,* 42, 683–708.

Brunnermeier, M. K., & Pedersen, L. H. (2008). Market liquidity and funding liquidity. *The Review of Financial Studies*, *22*(6), 2201-2238.

Cenedese, G., Sarno, L., & Tsiakas, I. (2014). Foreign exchange risk and the predictability of carry trade returns. *Journal of Banking & Finance*, *42*, 302-313.

Cochran, P. L., & Wood, R. A. (1984). Corporate social responsibility and financial performance. *Academy of Management Journal*, *27*(1), 42-56.

Datar, V. T., Naik, N. Y., & Radcliffe, R. (1998). Liquidity and stock returns: An alternative test. *Journal of Financial Markets*, *1*(2), 203-219.

Engle, R. F., & Ng, V. K. (1993). Measuring and testing the impact of news on volatility. *The Journal of Finance*, *48*(5), 1749-1778.

Escrig-Olmedo, E., Muñoz-Torres, M. J., & Fernández-Izquierdo, M. Á. (2013). Sustainable development and the financial system: society's perceptions about socially responsible investing. *Business Strategy and the Environment*, 22(6), 410-428.

Fama, E. F., & French, K. R. (2016). Dissecting anomalies with a five-factor model. *The Review of Financial Studies*, *29*(1), 69-103.

Fama, E. F., & MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, *81*(3), 607-636.

Fernando, C. S., Sharfman, M. P., & Uysal, V. B. (2009, August). Do investors want firms to be green? Environmental performance, ownership, and stock market liquidity. In *Academy of Management Proceedings* (Vol. 2009, No. 1, pp. 1-6). Academy of Management.

Friede, G., Busch, T., & Bassen, A. (2015). ESG and financial performance: aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance & Investment*, 5(4), 210-233.

Fu, F. (2009). Idiosyncratic risk and the cross-section of expected stock returns. *Journal of Financial Economics*, *91*(1), 24-37.

Giannarakis, G., & Theotokas, I. (2011). The effect of financial crisis in corporate social responsibility performance. International Journal of Marketing Studies, 3(1), 2.

Glosten, L.R., Jagannathan, R. & Runkle, D.E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance*, *48*(5), 1779-1801.

Gorton, G. & Metrick, A. (2012). Securitized banking and the run on repo. *Journal of Financial Economics*, *104*(3), 425-451.

Hayes, S. K. (2005). Socially responsible mutual funds: Issues to consider when investing with your conscience. *Journal of Financial Service Professionals*, *59*(5), 59.

Hill, R. P., Ainscough, T., Shank, T., & Manullang, D. (2007). Corporate social responsibility and socially responsible investing: A global perspective. *Journal of Business Ethics*, *70*(2), 165-174.

Hoberg, G. & Phillips, G. (2010). Product market synergies and competition in mergers and acquisitions: A text-based analysis. *The Review of Financial Studies*, *23*(10), 3773-3811.

Lean, H.H., Ang, W.R. and Smyth, R., 2015. Performance and performance persistence of socially responsible investment funds in Europe and North America. *The North American Journal of Economics and Finance*, *34*, pp.254-266.

Mancini, L., Ranaldo, A. & Wrampelmeyer, J. (2013). Liquidity in the foreign exchange market: Measurement, commonality, and risk premiums. *The Journal of Finance*, *68*(5), 1805-1841.

Moreira, A., & Muir, T. (2017). Volatility-Managed Portfolios. *The Journal of Finance*, *72*(4), 1611-1644.

Pástor, Ľ., & Stambaugh, R. F. (2003). Liquidity risk and expected stock returns. *Journal of Political Economy*, *111*(3), Stambaugh, 642-685.

Renneboog, L., Ter Horst, J., & Zhang, C. (2008). Socially responsible investments: Institutional aspects, performance, and investor behavior. *Journal of Banking & Finance*, *32*(9), 1723-1742.

Roca, E. (2013). The effect on price, liquidity and risk when stocks are added to and deleted from a sustainability index: Evidence from the Asia Pacific context. *Journal of Asian Economics*, *24*, 51-65.

Schueth, S. (2003). Socially responsible investing in the United States. *Journal of Business Ethics*, *43*(3), 189-194.

Shank, T., Manullang, D., & Hill, R. (2005). "Doing Well While Doing Good" Revisited: A Study of Socially Responsible Firms' Short-Term versus Long-term Performance. *Managerial Finance*, *31*(8), 33-46.

Souto, B. F. F. (2009). Crisis and corporate social responsibility: Threat or opportunity?. *International Journal of Economic Sciences and Applied Research,* 2(1), 36-50.

Van Liedekerke, L., De Moor, L. & Vanwalleghem, D. (April 2007). Risk-Return of Belgian SRI Funds. Available at SSRN: https://ssrn.com/abstract=1088979 or http://dx.doi.org/10.2139/ssrn.1088979

Performance Metrics for Credit Union Management

Robert B. Burney, Coastal Carolina University

Abstract

This paper discusses the potential weakness of the traditional financial institution performance ratios in assessing managerial performance for credit unions. Because the objective of credit unions is to redirect the gains from financial intermediation back to the credit union members, financial performance ratios designed for profit oriented financial institutions do not fully capture the effectiveness of credit union operations.

After discussing the basic issue and providing numerical examples, the paper moves on to discuss alternative complimentary measures which seek to address the weakness of the traditional financial institution performance metrics when used in the context of credit union performance assessment.

Performance Metrics for Credit Unions

As not-for-profit financial services cooperatives, credit unions provide some unique issues in evaluating performance. While regulators and managers continue to use traditional financial institution performance ratios to measure overall performance, such measures in their standard forms do not fully account for the actual mission of a financial cooperative. In particular, because a financial cooperative exists to serve its membership, a well-run credit union would not attempt to maximize profits, but would maximize value to the residual claimants – which in the cooperative form are the members. On a conceptual level, this is in keeping with the value maximization goal widely espoused in Anglosphere markets. But, how the created value is distributed is significantly different. Further, because financial cooperatives do not have tradeable shares, share price must be supplanted in the analysis by another global value creation measure.

When assessing managerial performance, many credit union boards use a balanced scorecard approach. A balanced scorecard approach typically includes key performance indicators which partially capture the return of value to members (ex. loans to shares), but more frequently are dominated by traditional financial performance ratios and metrics oriented toward institutional expansion.

The primary motivation for credit unions is to provide financial services to members in a format which transfers more of the overall value of the financial intermediation to the individual member and away from the institution. That is, good credit unions intentionally reduce their potential revenues by offering their members lower loan rates, while simultaneously intentionally increasing their cost of funds by offering their members higher deposit rates.

Operationally, this mission to benefit the credit union's members implies that good management should actually not attempt to maximize ROA and Net Interest Margin. Rather, good credit union management might instead attempt to judiciously push such measures down to the minimum level required by the institution's regulators. However, this minimization would need to occur subject to other constraints which would ensure that low ratios were not simply resulting from poor managerial capabilities. Thus, some measure of the loss in traditional performance metrics attributable to increased member benefit is needed. Before describing such a framework, it is useful to review the dominant framework for analyzing the financial performance of financial institutions.

Analytical Framework for Financial Institutions

A common approach to analyzing the financial performance of financial institutions is derived from the operational dynamic reflected by the income statement format used by financial institutions. The financial institution income statement reflects the business model wherein interest revenues less interest expense is the main driver of profit with net non-interest expense (operating expenses less fees), provision for loan losses, and taxes being the additional key elements. We begin by writing the financial institution income statement horizontally as follows.

$$Net\ Income = Net\ Interest\ Income - Net\ Noninterest\ expense - provision\ for\ loan\ losses - taxes \qquad (1)$$

For credit unions, which are not-for-profit, taxes equal zero. So, equation (1) changes as follows.

$$Net\ Income = Net\ Interest\ Income - Net\ Noninterest\ expense - Provision\ for\ loan\ losses \qquad (2)$$

Next, both sides of the equation are scaled by total assets, which results in the following.

$$\frac{Net\ Income}{Total\ Assets} = \frac{Net\ Interest\ Income}{Total\ Assets} - \frac{Net\ Noninterest\ expense}{Total\ Assets} - \frac{Provision\ for\ loan\ losses}{Total\ Assets} \qquad (3)$$

At this point we arrive at a common deconstruction of ROA for financial institutions. Net Interest Income divided by Total Assets is Net Interest Margin (NIM). Net Noninterest Expense divided by Total Assets is the "burden" ratio. And, Provision for Loan Losses divided by Total Assets is simply %PLL. Thus, we have the following.

$$ROA = NIM - Burden\ Ratio - \%PLL \qquad (4)$$

We can use this framework to discuss the complicating factor of credit unions' cooperative format in the context of financial institutions managerial performance measurement. While it may be intuitive to suggest that a credit union should strive toward the lowest, safe level of ROA in order to ensure that maximum value is distributed to its membership, the deconstructed elements present more insight, and potential measurement.

First, let us consider Net Interest Margin. Net Interest Margin is a measure of the fundamental earning potential of the financial institution. Net Interest Income is defined as Interest Income minus Interest Expense. In general, financial institutions would be expected to attempt to maximize this ratio by increasing Interest Income when possible while decreasing Interest Expense when possible.

$$NIM = \frac{(Interest\ Revenue - Interest\ Expense)}{Total\ Assets} \qquad (5)$$

Interest Income would be maximized by charging the highest rates on loans which the market would bear, given some target level of lending. Interest Expense would be minimized by paying the lowest deposit rates possible while still attracting an adequate amount of deposits. But, for credit unions, the net impact of such actions would be to cause increased net income to accumulate as undistributed earnings. While for profit financial institutions distribute a portion of profits as dividends, credit unions consider the interest paid on interest bearing accounts to be dividends.

Performance Measurement Implications for Credit Unions

If we consider the merit in distributing more of the value arising from the financial intermediation process to a credit union's members, then a lower NIM would be preferred to a higher NIM, provided the lower NIM were still at a level expected to build an adequate net worth cushion against losses – and, that the funds of the institution were being put to good use through an appropriate level of lending (loan-to-share ratio).

Here on the interest income side we would want to differentiate between charging low member rates on a large number of member loans and simply not making many loans. The first is an admirable credit union goal, while the second suggests unmotivated management. On the interest expense sides, we would want to differentiate between paying high deposit rates on a large number of member deposits and simply paying far too much on deposits. One way to make this assessment it to compare what NIM would have been had the credit union charged higher feasible market loan rates and paid lower feasible market deposit rates. Conceptually, the metric would be based on the following differential, were NIM would be the reported value and NIM' would be the ratio's value which would have resulted if the institution had accrued the potential value of financial intermediation instead of intentionally distributing it to the credit union members. This is set forth in equation (6) and expanded in equations (7) and (8).

$$Management\ Differential = NIM' - NIM \qquad (6)$$

$$Management\ Differential_{NIM} = \frac{(Interest\ Revenue' - Interest\ Expense')}{Total\ Assets} - \frac{(Interest\ Revenue - Interest\ Expense)}{Total\ Assets} \qquad (7)$$

$$Management\ Differential_{NIM} = \frac{(Interest\ Revenue' - Interest\ Revenue) - (Interest\ Expense' - Interest\ Expense)}{Total\ Assets} \qquad (8)$$

As a simple numerical example, consider a credit union with $100 mil. in loans, $10 mil. in cash and fixed assets, $100 mil. in deposits, and $10 in net worth. Further assume that the member focused credit union has an average rate of 5% on its loans and pays an average rate of 3% on its deposits. Given that most financial institutions conduct or have access to ongoing market interest rate comparison studies, we could also know, that based on reported competitor rates, a non-member focused institution could have charged an average rate of 6% on its loans and paid only 2% on its deposits.

In this example the actual NIM reported would be as follows in Eqn. 9, while the potential NIM' would be as in Eqn. 10.

$$NIM = \frac{(\$5\ mil. - \$3\ mil.)}{\$110\ mil.} = .018 \tag{9}$$

$$NIM = \frac{(\$6\ mil. - \$2\ mil.)}{\$110\ mil.} = .036 \tag{10}$$

From this perspective, the managers have given a potential .018 return on assets away to the membership. In dollar terms, the additional value delivered to the membership is $2 mil. in this example. However, it is important to consider the volume of loans and deposits is identical in this example. A more detailed approach might attempt to estimate the decrease in volume of loans and deposits from higher loan rates and lower deposit rates. In either case, the implication is that a manager's performance comparison to a peer group benchmark can be better understood with this added dimension.

Next let us turn our attention to the burden ratio. The burden ratio measures the institution's non-interest operating expenses as reduced (or offset) by the amount of non-interest income (predominantly fees). Here too, the credit union mission could lead to an apparent inferior performance. The burden ratio is defined in more detail in Eqn. 11.

$$Burden\ Ratio = \frac{Noninterest\ expense - Noninterest\ income}{Total\ Assets} \tag{11}$$

With the burden ratio, the adjustment would be slightly more straightforward since noninterest expense should be minimized regardless of the profit or not-for-profit status of the institution. For example employee and occupancy expense would not be expected to be increased simple because of the cooperative format. However, it would be reasonable to assume that noninterest income (fees to the membership) would be minimized. Thus an adjustment similar to that done above for NIM is possible. Here again, benchmarking competitors' fees would be required.

The net noninterest expense construct has operating expenses (ex. occupancy, employees, utilities) being offset or reduced by noninterest income (ex. fees). The management of a member focused credit union attempting to return financial intermediation value to members would be expected to attempt to minimize operating expenses. But, the same management would be expected to minimize fees charged to members. Thus, management would tend intentionally to increase the institution's burden ratio. In assessing managerial performance, the approach would be as presented in eqn. 12 below.

$$Management\ Differential_{BR} = BR - BR' \tag{12}$$

Note that the differential here is in net operating cost. Thus, the value-to-members impact would be the surplus of current burden ratio over the potential burden ratio. That is, the higher net non-interest expense resulting from the lower offset from charging lower fees to members is directly related to the value transfer to members. Adding detail, the relationship is as follows.

$$\begin{aligned} &Management\ Differential_{BR} \\ &= \frac{Noninterest\ expense - Noninterest\ income}{Total\ Assets} - \frac{Noninterest\ expense' - Noninterest\ income'}{Total\ Assets} \end{aligned} \tag{13}$$

Assuming that operating expense should not be impacted by the variation in managerial objective, we would have the following. So, the additional value distributed to members would represent a foregone portion of potential ROA.

$$Management\ Differential_{BR} = \frac{Noninterest\ income' - Noninterest\ income}{Total\ Assets} \tag{14}$$

Assuming the same numerical values from the hypothetical example institution above, suppose that the member-oriented credit union might charge members only $ .4 mil. fees, whereas the market would bear $.7 mil. in fees for the same transaction patterns. Thus, the value-to-members objective would transfer $.3 mil. in value to the members, but would reduce ROA by .003.

Because financial institutions typically monitor competitor's fee structure, this calculation should not be too difficult. However, as before in the case of NIM, the comparison would depend on a comparable level of transaction volume. So, the impact of changing fees on the number of transactions would need to be considered in certain cases.

Finally, some mention must be made of the %PLL in the deconstruction given in equation (4). The provision for loan losses is a charge against current income which is an estimate of the loss associated with bad debts. For financial institutions, this

charge accumulates as a contra-asset balance sheet item called Allowance for Loan Losses. This item is then reduced when actual loan charge offs occur.

The provision for loan losses is calculated under standardized approaches, so it is interpreted as being a measure of loan quality. The higher the credit quality of a loan portfolio, the lower the associated provision for loan losses should be over time. Consequently, %PPL is a measure of the quality of the loan portfolio.

Our discussion above largely assumes that the default characteristics of the loan portfolio held by the institution would not be impacted by the actual organizational form. Thus, the managerial objective of maintaining a high-quality loan portfolio would not differ between credit unions and for-profit financial institutions.

However, one perspective suggests that credit unions should also aspire to provide financial services to those who might otherwise not be able to access financial services. For those credit unions for which this is an important operational goal, the use of the alternative metrics discussed above would be further complicated since the rates offered on loans would need to be further adjusted to account for differences in loan portfolio quality. Data required for that adjustment would include not just rate differentials overall, but credit spreads between different loan quality classes.

Implementation of Alternative Performance Metrics

As mentioned at the outset, these types of proposed alternative management performance metrics are envisioned as part of a multifaceted balanced score card assessment. Such alternative metrics would augment commonly used performance ratios. Because they would provide a measure of performance which explicitly accounted for the organizational structure of the credit union, these types of ratios could assist board members in gaining a deeper understanding of managerial performance.

However, because these are nonstandard measures, peer group comparisons would likely be more difficult. A given credit union could fairly easily calculate its own value for these proposed metrics since many would already have the required inputs (competitor's rates and fees). This would also make it relatively easy to calculate the value of these metrics for any other credit union operating in the same local market.

Comparison to other credit union peers is complicated by the assumption that all credit unions operate based on similar motivation. And, credit union benchmarking is usually done based on peer group definitions primarily driven by size and only secondarily by operating characteristics and geographical location. Thus, the greatest value in this type of comparison would be to community (non-SEG) credit unions operating in a competitive environment in which similarly sized for-profit financial institutions were strong credit union competitors.

For credit unions operating in a competitive community bank environment, the additional dimension is particularly appealing. This is the case because such community credit unions frequently operate in a fashion much more similar to community banks than a large single SEG credit union. Competing in the community banking space typically means a larger branch network and potentially replicating functions (ex. business lending) which are required of community financial institutions.

Summary

This paper has discussed the potential for improving credit union management by adding nonstandard performance metrics to a balanced scorecard approach. The proposed alternative measures are intended to measure managers' success with respect to the fundamental credit union goal of transferring the value arising from the financial intermediation process to credit union members. Because of differences in how for-profit and not-for-profit financial institutions transfer this value to the residual claimant group, for-profit financial performance ratios may not fully capture this performance when applied to not-for-profit credit unions.

In particular, the residual claimants (shareholders) of a for-profit financial institution receive this value in the form of dividends and increases in the (liquidatable) market value of their shares. In contrast, the residual claimants (members) of a not-for-profit credit union receive this value in the form of dividends on deposits, but also from intentional reductions in profitability arising from lower revenues (lower loan rates) and higher costs (lower cost-offsetting fees). Consequently, the for-profit based ratios potentially fail to capture the level of managerial performance.

This nuanced issue requires that we differentiate between simple poor managerial performance at credit unions, and good managerial performance which intentionally reduces the potential values of the common metrics. To this end, the proposed auxiliary performance metrics presented in this paper draw quantitative contrasts between the credit union's values of the standard ratios and the potential value which could have resulted if for-profit motives would have been in place.

Certainly, credit union regulators are not likely to deviate from their current use of the standard financial performance metrics. The regulators are primarily interested in safety and soundness, and define their overall ratings (ex. CAMEL ratings) with respect to those goals. Indeed, the regulators assessments are unidirectional. It is not common for a credit union regulator

to criticize management for having excessively high values for ROA, NIM, and Net Worth. However, internally, a credit union's board of directors can use the proposed alternative measure to assess management's performance subject to the regulator constraint.

References

Callahan and Associates. (DL2018). *The Comprehensive Guide to Credit Union Performance Benchmarking.* Callahan.com

Chesnick, David S. 2000. *Financial Management and Ratio Analysis For Cooperative Enterprises*. USDA Rural Business–Cooperative Service Report 175.

Chiou, Wan-Jiun, Robert L. Porter, and Ben J. Sopranzetti. 2014. "Understanding the Channels of Bank Value Creation During Times of Crisis: Deconstructing ROA". *Journal of Accounting and Finance* vol. 14(5), pp. 11-23.

CMPG, LLC. 2015. *The Importance of the Burden Ratio*. 2015 Intelligent Bank Management Series.

Gardner, Mona J., Dixie L. Mills, and Elizabeth S. Cooperman. 2005. *Managing Financial Institutions*. Boston, Cengage.

Higgins, Mike, Jr. and Dennis W. Campbell. 2012. *Improving Peer Group Analysis for Credit Unions*. Madison, Wisconsin, Filene Research Institute.

Simkhada, Nav Raj. 2017. "Indicators for Measuring Performance of Financial Cooperatives in Nepal". *Journal of Business and Management Research*, (July), Vol.2, No.1 & 2, pp. 66-86

Financial Literacy After-School Programs: An Exploratory Study

Nathan Castro, Jennifer Brodmann, California State University Dominguez Hills

Abstract

We examine the impact of financial literacy in after-school programs to determine the structure needed in order to improve financial literacy rates in adolescents and adults. We have also recognized the need for financial literacy with students preparing to be financially independent. We conduct a survey of the literature on the types of financial literacy programs offered in the United States, their program outcomes, and what can be improved upon to increase efficacy. From these findings, we propose a structure and curriculum for after-school programs that will give students the skills to improve their financial decision-making.

Introduction

The lack of financial literacy in the United States is a significant issue that needs to be addressed. Financial literacy is a major driver of wealth inequality (Lusardi, Michaud, and Mitchell, 2013). Financial literacy is necessary for individuals to address both short-term and long-term needs. Poor financial choices can put individuals and households into debt and leave limited if no resources for major expenses, such as education, homeownership, healthcare, retirement, and unexpected emergency costs. These individuals may also be left susceptible to scams or predatory lending. Many households also fall behind on maintaining a household budget, which leaves households exposed to higher risks of financial distress. Not only will these financial decisions affect themselves, but their future generations from building intergenerational wealth.

One of the most vital predictors of financial literacy is socioeconomic status. Students need equal access to financial resources to address gaps in financial education. If not, this issue can persist and become more of an economic disparity later in life and affect their family's financial well-being for future generations. This exposes the need for financial literacy programs that serve disadvantaged students and address wealth inequality.

This study provides an overview of the recent literature on financial literacy and financial literacy programs and proposes a new framework for after-school programs to address financial literacy. We examine previous studies on financial literacy programs and assessments for success and create a new framework for financial literacy programs that address the issues faced in previous programs. This study is different than others because it combines the utilization of afterschool programs with preparation for financial independence. Beverly and Burkhalter (2005) discuss the importance of financial literacy and propose that financial education should be incorporated into public school curriculums, and is an important topic needing acknowledgment from social workers. Beverly and Burkhalter (2005) emphasize that youths exhibit low levels of financial literacy and several do not engage in responsible financial practices, such as budgeting and utilizing credit in a discerningly. We survey the literature on financial literacy and propose an after-school curriculum for high-school, college students, and adults to learn financial literacy to make more informed financial decisions.

We have recognized the need for financial literacy for students. When students start to get jobs and graduate, most are not equipped with enough financial intelligence to be self-sustained financially. It was discovered that afterschool programs help equip students for further traditional education, such as STEM programs, according to Karp and Maloney (2013). Our research suggests that there is a need for financial education with those who are preparing to become financially independent.

Literature Review

The Importance of Financial Literacy

Various studies provide valuable information regarding the contemporary structure of financial literacy. Braunstein and Welch (2002) provide an overview of the current research on financial literacy and its practices and policies. Beverly and Burkhalter (2005) note the consequences of those who have not received financial education are unprepared when financial hardships arise. Situations of rising debt, or losing out on opportunities such as a mortgage loan are some examples of what can happen without financial literacy. Financial literacy can be related to household saving. Deuflhard, Georgarakos, and Inderst (2015) find that more financial literate investors receive higher savings account returns. In order to successfully educate students on finance, a structure and curriculum must be in practice. Donovan, Weiner, Gross, and Block-Lieb (2005) evaluate

a two-day financial literacy teaching training program. This training equipped teachers with skills that improve their efficiency. Trained teachers resulted in more confidence and being better prepared for class engagement.

Ward (2013) believes that families need to start worrying about their children's finances as soon as possible. Making financial literacy a focus early on helps parents and children prepare for savings. This will help lessen or eliminate the financial burden of student loans (Ward, 2013). Lusardi and Mitchell (2014) acknowledge the recent growth of research about financial literacy and note that being financially ignorant can result in a voluntary payment of fees that could have been avoided through passive indexing. Financial literacy is a necessity because it has a direct association with net worth (Rooij, Lusardi, and Alessie, 2012). This relationship is important because net worth is a key variable in the determination of wealth (Rooij, Lusardi, and Alessie, 2012). Calcagno and Monticone (2015) suggest that those without financial literacy are prone to making poor financial decisions. Examples of poor decision making would include a lack of preparation for emergencies or retirement. Those without financial literacy can still find solace in financial advisors (Monticone, 2015).

Financial Literacy Education

MainSource Bank (2016) recognized the growing need for children with financial literacy. MainSource Bank designed a SuperSaver account that was made to teach financial responsibility while establishing a savings account. There is no minimum balance and no monthly fee, so children are not intimidated when initially learning about their finances. Financial advisors suggest that some of the best financial advice is not exceeding your means, paying yourself first, and budgeting responsibly (Benzinga, 2012). Parents regularly use their personal experience with money as the only means of reference when educating their children on finance (Benzinga, 2012). This is important; however, the content used is limited to the experience of close relatives.

Types of Financial Literacy Programs

There are several types of financial literacy programs and initiatives that have been implemented to address this lack of financial literacy. We will discuss these different types of programs as well as the literature that has assessed their effectiveness. GreenPath Debt Solutions (2015) is an educational project made to bring financial literacy programs to public libraries.

Financial Literacy and Education Programs

According to a recent amendment to the U.S. Bankruptcy Code, certain individual debtors who file chapter 7 and 13 bankruptcy cases are required to complete a financial literacy course. Baron-Donovan, Wiener, Gross, and Block-Lieb (2005, pp. 65) study the teacher training for this financial literacy course, which consisted of five two-day training sessions, and find that questionnaires pre- and post-test reflect a 9 percent increase in financial knowledge. The training covered three main areas, which are "1) encouraging debtors to examine the psychological underpinnings of spending decisions; 2) helping consumer debtors develop spending and savings plans; 3) providing information that would enable debtors to receive the maximum benefit from the bankruptcy system's fresh start; 4) making individual debtors more alert to uses and misuses of credit; and 5) providing debtors with a vocabulary so that they would be better able to make thoughtful financial decisions."

Pang (2009) reported on a study that administers and assesses the results of financial literacy programs for adolescents. The specific study reported on was comprised of two lessons conducted by 12 teachers with an average of 10 years teaching each (Pang, 2009). The two lessons included a learning study group and a lesson study group (Pang, 2009). 16 40 minute lessons were given with a multiple-choice pre-test and the three post-tests. The first post-test was given immediately after the study groups finish their lessons. The second post-test was given six weeks after the study groups. The last post-test was given six months after the study group. The only real difference between the two groups was that the lesson study group was administered by the use of Variation Theory (Pang, 2009). This theory suggests that it is necessary to provide subjects with differences when teaching financial literacy (Pang, 2009).

Homebuyer Counseling Programs

Several affordable housing programs that serve prospective homeowners that have low income, low credit, and limited savings provide financial literacy training to teach skills in debt management, budgeting, and saving. This strategy of financial literacy training for prospective homebuyers has been implemented by several community organizations, with the Neighborhood Housing Services, which is a subsidiary of the Neighborhood Reinvestment Corporation, being one of the first community organizations that implemented this training practice.

Savings Initiatives

Financial literacy training is a major part of initiatives to promote lower- and middle-income household saving. In 2001, The Consumer Federation of America launched America Saves, a program where communities promote local savings campaigns. This program included enrolling residents as savers, offer them no-fee savings accounts, offer motivational workshops, as well as consulting one-on-one. An area-wide survey was conducted and suggested that 10,000 Cleveland-area residents were motivated to become more effective in saving. This program has expanded to over 60 local campaigns offered throughout the U.S.[i] The program Money 2000, sponsored by the U.S. Department of Agriculture, was launched to offer information and the tools needed for consumers to improve their saving habits and become more responsible with spending. The participants of this program increased their on average savings by $1,600 over twelve-months and reduced their average credit balances b by over $1,200.

Matched savings programs, also called individual development accounts (IDAs), were created to help lower-income households who did not have access to employer-sponsored savings programs. The program participants would open a savings account were required to have a specific saving objective with their contributions being matched by sponsoring organizations. These matching funds would be forfeited if the participant withdrew the money for any other reason besides buying a home, opening a business, or paying for higher education. The Corporation for Enterprise Development (CED) assessed the effectiveness of the IDAs. The participants went through 10.5 hours of financial training and their savings were reported from September 1997 until June 2000. The CED found that financial training had an essential role in participant saving, with their assessment revealing that an increase of the average monthly net deposit resulted in each additional hour of training of up to twelve hours.

Workplace Programs

Due to the shift retirement plans being employer-driven to employee-driven, many employees need to assume greater financial responsibility in their retirement planning. Because of this, employers have implemented training seminars to help employees assess their needs and plan for retirement. Employee response to these programs has primarily been positive. These types of programs have been found to help employee increase their financial well-being (Garman, Kim, Kratzer, Brunson, and Joo, 1999), increase participation in 401(k) plans (Kim, Constance, Kratzer, and Leech, 2001), and was effective in adjusting their financial behavior (Kim, 2001).

Main Findings

From the previous literature, we find that financial literacy programs are needed to better prepare individuals for making informed financial decisions. The major areas that individuals face are how to utilize credit, save money for unexpected costs (such as illness), and plan for retirement (Brazel, Jones, Thayer, and Warne, 2015). We also found that planning for future financial situations such as college tuition can lessen or eliminate the need to take out student loans. The 2008 recession was used as an example of what could go wrong when the population is not financially literate (Brazel, Jones, Thayer, and Warne, 2015). Luckily, there has been a recent surge in those who notice the demand for financial literacy (Brazel, Jones, Thayer, and Warne, 2015).

Resources of programs or savings initiatives for adolescents have become a focal point of contemporary banks such as MainSource (Rooij, Lusardi, and Alessie, 2012). Previous literature shows us the relationship between financial literacy and net worth while recognizing the need for financial education. Developing financial knowledge is positively related to a boost in wealth when retirement planning (Rooij, Lusardi, and Alessie, 2012).

Afterschool programs that focus on specified subjects such as STEM have shown to be successful (Karp and Maloney, 2013). There is a positive correlation between program engagement and student participation in class. The benefits of these programs extend beyond the students and teachers. College-level students and volunteers can also benefit from experience or internships. While there are afterschool programs and financial literacy programs, there are minimal opportunities for students to receive financial education within an afterschool program.

Outcomes

Beverly and Burkhalter (2005) conclude that contemporary youth are not likely to engage in financial practices, resulting in lower levels of financial literacy. There was a positive correlation between financial counseling and a reduction in debt and account usage (Journal of Consumer Affairs, 2007). Financial literacy is an essential skill that needs to be development (Lusardi, 2015). Afterschool programs targeted towards specified subjects have shown to be successful (Karp and Malone,

2013). The success of similar programs and the demand for financial literacy create an ideal environment for an afterschool program specified for financial education.

Employer retirement pension programs are now becoming a normality (Lusardi, 2008). Employees are automatically rolled into pension plans where employees have the option to opt-out or not. Set contribution rates of the employer incentivize these programs. These programs have shown success with participation rates primarily due to the employer contribution incentives. Lusardi (2008) acknowledges the need for these programs by pointing out the lack of financial literacy adolescence and the elderly. Programs like these help employees develop financial literacy without having to resort to third-party programs. This is helpful because it reflects similar aspects of what in after-school financial literacy program would include.

Pang (2009) conclude that their learning study group was 5% more effective in their six-month post-test than the lesson study group. This shows that effective communication of financial concepts such as inflation, savings, opportunity costs, and retirement can result in a growth in financial literacy. More importantly, this study shows that there are various approaches to teaching financial literacy to adolescence. Deciphering which methods are more successful lays a foundation for future financial teaching practices to adolescents.

Areas of Improvement

The Sharpen Your Financial Focus program by the National Foundation for Credit Counseling studied by researchers from Ohio State University. Class participants were required to take a quiz to determine their financial skills in budgeting saving, managing debt, and preparing for retirement. The participant reviewed the results with a counselor, and the counselor would identify which areas needed improvement and would target the education to those specific needs through in-person workshops and one-on-one financial coaching (Weston, 2016). From this study, the researchers created a control group and when comparing the group with the group who went through the counseling program, the researchers found that their clients made a significant reduction in their revolving debt by over $2700 and their overall debt by over $7600. Their clients reported that they felt more confident in managing their finances and paying their debts on a more consistent basis. This process of setting goals, targeting specific education workshops, and one-on-one coaching resulted in positive financial outcomes for these clients. Having programs adhere to national standards can also increase the effectiveness of the programs. The Council for Economic Education provides national standards on want concepts students should know at each grade level and provides curriculum support[2].

Proposed Framework for a Financial Literacy After-school Program

From examining the previous literature, we propose the development of financial literacy programs that contain the following components:
- Adhere to national financial literacy standards from The Council for Economic Education
- Have students take an assessment test prior to going through the program to understand what financial skills the student has and how to build upon them and then have them take an exit test to see where their skills are after going through the program
- Provide financial coaching that helps students create specific goals and routinely follow-up on the progress of attaining those goals
- Provide a targeted list of the workshops that the student should take to build specific needed skills
- Offer both in-class, small-group, and alternative media instruction for students based on student learning preference

Conclusion and Future Research

Our study provides an overview of the research literature on financial literacy programs. We discuss the importance of financial literacy in increasing the financial skills needed for individuals so they can make more responsible financial decisions.We also acknowledge the effect that financial literacy has in preparation for becoming financially independent. Equipping students with financial tools can help hedge against going into a significant amount of debt before graduating. We find that afterschool programs focused on a specified subject creates engagement. The demand for financial literacy programs is growing, and establishing an afterschool program creates an incentivized convenience for students to plan for their future.

We will expand our research and conduct a survey on adults to determine their interest in themselves and their children taking financial literacy classes. We will administer surveys and conduct focus groups in community centers on assessing the needs to serve the population with financial education that will increase financial literacy and responsible financial decision-making. This may have a significant impact on the community by equipping adolescents and adults with knowledge on how to

take financial responsibility. With overall rising student debt, and a growing concern of finances (Braunstein and Welch, 2002), an afterschool program educating students in financial literacy is needed. This may be an engaging and practical way to prepare students for financial independence.

Acknowledgment

Notes

1. https://americasaves.org/for-savers/connect-to-a-campaign
2. https://www.councilforeconed.org/programs-2/curriculum-support/

References

Arthur, Chris. "Financial literacy education." *Financial Literacy Education*. SensePublishers, Rotterdam, 2012. 1-12.

Baron-Donovan, C., Wiener, R. L., Gross, K., & Block-Lieb, S. (2005). Financial literacy teacher training: A multiple-measure evaluation.

Benzinga (2012). Servicemembers Focus on Financial Literacy For Children, First Command Reports. *Benzinga*. Retrieved October 4, 2018.

Beverly, S. G., & Burkhalter, E. K. (2005). Improving the financial literacy and practices of youths. *Children & Schools*, *27*(2), 121.

Braunstein, S., & Welch, C. (2002). Financial literacy: An overview of practice, research, and policy. *Fed. Res. Bull.*, *88*, 445.

Brazel, J. F., Jones, K. L., Thayer, J., & Warne, R. C. (2015). Understanding investor perceptions of financial statement fraud and their use of red flags: Evidence from the field. *Review of Accounting Studies, 20*(4), 1373-1406. Retrieved October 21, 2018.

Brooks, G., Burton, M., Cole, P., Miles, J., Torgerson, C., & Torgerson, D. (2008). Randomised controlled trial of incentives to improve attendance at adult literacy classes. *Oxford Review of Education*, *34*(5), 493-504.

Bucher-Koenen, T., & Ziegelmeyer, M. (2013). Once burned, twice shy? Financial literacy and wealth losses during the financial crisis. *Review of Finance*, *18*(6), 2215-2246.

Calcagno, R., & Monticone, C. (2015). Financial literacy and the demand for financial advice. *Journal of Banking & Finance, 50*, 363-380. Retrieved October 5, 2018.

Carlin, B. I., & Robinson, D. T. (2012). What does financial literacy training teach us?. *The Journal of Economic Education*, *43*(3), 235-247.

Cude, B., Lawrence, F., Lyons, A., Metzger, K., LeJeune, E., Marks, L., & Machtmes, K. (2006). College students and financial literacy: What they know and what we need to learn. *Proceedings of the Eastern Family Economics and Resource Management Association*, *102*(9), 106-109.

Deuflhard, F., Georgarakos, D., & Inderst, R., (2015). 'Financial literacy and savings account returns,' IMFS Working Paper Series 88, Goethe University Frankfurt, Institute for Monetary and Financial Stability (IMFS).

Drexler, A., Fischer, G., & Schoar, A. (2014). Keeping it simple: Financial literacy and rules of thumb. *American Economic Journal: Applied Economics*, *6*(2), 1-31.

Faulkner, A. E. (2017). Financial literacy education in the United States: Exploring popular personal finance literature. *Journal of Librarianship and Information Science*, *49*(3), 287-298.

Friedman, P. (2005). Providing and Funding Financial Literacy Programs for Low Income Adults and Youth. Strategy Brief. *Finance Project*.

Gale, W., & Levine, R. (2010). Financial Literacy: What Works? How Could It Be More Effective?.

Garman, E. T., Kim, J., Kratzer, C. Y., Brunson, B. H., & Joo, S. H. (1999). Workplace financial education improves personal financial wellness. *Financial Counseling and Planning*, *10*(1), 79-88.

GreenPath, Consumer Financial Protection Bureau, National Foundation for Credit Counseling offer personal finance library programs. (2015). *M2 Banking & Credit News (BCN)*. Retrieved October 5, 2018.

Grohmann, A., Kouwenberg, R., & Menkhoff, L. (2015). Childhood roots of financial literacy. *Journal of Economic Psychology*, *51*, 114-133.

Guiso, L., & Viviano, E. (2014). How much can financial literacy help?. *Review of Finance*, *19*(4), 1347-1382.

Hastings, J. S., Madrian, B. C., & Skimmyhorn, W. L. (2013). Financial literacy, financial education, and economic outcomes.

Jacob, K., Hudson, S., & Bush, M. (2000). Tools for Survival: An Analysis of Financial Literacy Programs. *Chicago: Woodstock Institute*.

Johnson, E., & Sherraden, M. S. (2007). From Financial Literacy to Financial Capability among Youth. *J. Soc. & Soc. Welfare*, *34*, 119.

Jorgensen, B. L. (2007). *Financial literacy of college students: Parental and peer influences* (Doctoral dissertation, Virginia Tech).

Karp, T., & Maloney, P. (2013). Exciting Young Students In Grades K-8 About STEM Through An Afterschool Robotics Challenge. *American Journal of Engineering Education (AJEE), 4*(1), 39.

Kefela, G. T. (2010). Promoting access to finance by empowering consumers-Financial literacy in developing countries. *Educational Research and Reviews*, *5*(5), 205-212.

Kezar, A., & Yang, H. (2010). The Importance of Financial Literacy. *Wiley InterScience, 14*(6), 15-21. Retrieved October 5, 2018.

Kim, J., Kratzer, C, Leech I. (2005) 'Impacts of Workplace Financial Education on Retirement Plans' in Jeanne M. Hogarth ed. Proceedings of the 1002 Annual Conference of the Association for Financial Counseling and Planning Education, 28.

Kim, J. 'The Effectiveness of Individual Counseling Financial Counseling Advice' in Jeanne M. Hogarth ed. Proceedings of the 1002 Annual Conference of the Association for Financial Counseling and Planning Education, 62-69.

Korutaro Nkundabanyanga, S., Kasozi, D., Nalukenge, I., & Tauringana, V. (2014). Lending terms, financial literacy and formal credit accessibility. *International Journal of Social Economics*, *41*(5), 342-361.

Ludlum, M., Tilker, K., Ritter, D., Cowart, T. W., Xu, W., & Smith, B. C. (2012). Financial literacy and credit cards: A multi-campus survey.

Lusardi, A. (2015). Financial literacy skills for the 21st century: Evidence from PISA. *Journal of Consumer Affairs*, *49*(3), 639-659.

Lusardi, A. (2008). *Financial literacy: an essential tool for informed consumer choice?* (No. w14084). National Bureau of Economic Research.

Lusardi, A., & Mitchell, O. S. (2014). The economic importance of financial literacy: Theory and evidence. *Journal of Economic Literature*, *52*(1), 5-44.

Lusardi, A., Michaud, P., & Mitchell, O. (2017). Optimal Financial Knowledge and Wealth Inequality. *Journal of Political Economy, 125*(21). Retrieved October 21, 2018.

Lusardi, A. (2008). Overcoming the Saving Slump: How to Increase the Effectiveness of Financial Education and Saving Programs. *Chicago: University of Chicago Press.* Retrieved October 12, 2018.

MainSource Bank Focuses Strong Attention on Financial Literacy for Children. (2016). *Plus Media Updates*. Retrieved October 4, 2018.

Mandell, L. (2006, October). Teaching young dogs old tricks: The effectiveness of financial literacy intervention in pre-high school grades. In *Academy of Financial Services 2006 Annual Conference, Salt Lake City, UT*.

Mandell, L. (2006). Financial Literacy: If It's So Important, Why Isn't It Improving?.

Mandell, L. (2008). The financial literacy of young American adults. *The Jumpstart Coalition for Personal Financial Literacy*.

Mandell, L., & Klein, L. (2009). The impact of financial literacy education on subsequent financial behavior.

Mandell, L., & Klein, L. S. (2007). Motivation and financial literacy. *Financial services review*, *16*(2), 105.

Mandell, L., & Hanson, K. O. (2009, January). The impact of financial education in high school and college on financial literacy and subsequent financial decision making. In *American Economic Association Meetings* (Vol. 4, p. 2009).

Meier, S., & Sprenger, C. D. (2013). Discounting financial literacy: Time preferences and participation in financial education programs. *Journal of Economic Behavior & Organization*, *95*, 159-174.

Moulton, S., & Roll, S. (2016). The NFCC® Sharpen Your Financial Focus Initiative Impact Evaluation. Retrieved October 4, 2018, from https://nfcc.org/wp-content/uploads/2017/05/134236_NFCC_WEB.pdf.

Norvilitis, J. M., Merwin, M. M., Osberg, T. M., Roehling, P. V., Young, P., & Kamas, M. M. (2006). Personality factors, money attitudes, financial knowledge, and credit-card debt in college students 1. *Journal of Applied Social Psychology*, *36*(6), 1395-1413.

Pang, M. F. (2010). Boosting financial literacy: Benefits from learning study. *Instructional Science*, *38*(6), 659-677.

Reich, C. M., & Berman, J. S. (2015). Do financial literacy classes help? An experimental assessment in a low-income population. *Journal of Social Service Research*, *41*(2), 193-203.

Rooij, M. C., Lusardi, A., & Alessie, R. J. (2012). *The Economic Journal, 122*(560). Retrieved October 5, 2018.

Rutledge, S. L. (2010). *Consumer protection and financial literacy: Lessons from nine country studies*. The World Bank.

Sprow Forté, K. (2013). Educating for financial literacy: A case study with a sociocultural lens. *Adult Education Quarterly*, *63*(3), 215-235.

The impact of credit counseling on subsequent borrower behavior. (2007). *Journal of Consumer Affairs*. Retrieved October 5, 2018.

Van Rooij, M. C., Lusardi, A., & Alessie, R. J. (2012). Financial literacy, retirement planning and household wealth. *The Economic Journal*, *122*(560), 449-478.

Ward, L. (2013). Investing in Funds & ETFs: A Monthly Analysis --- Family Finance: Save for Yourself -- or Kids? --- It's a juggling act for families facing both retirement and college costs. *Wall Street Journal*. Retrieved October 4, 2018.

Weston, L. (2016, September 12). Liz Weston: Money lessons that actually stick. Retrieved October 4, 2018, from https://www.dailyherald.com/article/20160912/business/309129905/

Willis, L. E. (2009). Evidence and ideology in assessing the effectiveness of financial literacy education. *San Diego L. Rev.*, *46*, 415.

Does Investing In Private Shares Make Sense For Mutual Funds?

Marlin R.H. Jensen and John S. Jahera, Jr., Auburn University

Abstract

With the rise of index funds, actively managed funds are looking for ways to distinguish themselves by buying into private companies. Mutual funds can invest up to 15 percent of their assets in restricted securities. Using the Fidelity's flagship Magellan Fund, we explore how Magellan has accounted for their private investments performance and whether Magellan has manipulated its performance. In addition, we analyze the performance of the private holdings through time to assess the recorded performance prior to the private holding going public or being sold and the subsequent performance of the holdings after the private companies have gone public.

Introduction

Mutual funds have become an investors' vehicle of choice because of flexibility and transparency they provide. Holdings of a fund are fairly valued on a daily basis and investors can contribute or withdraw money each day. Therefore, mutual funds need to follow certain guidelines. Holdings of a fund must have sufficient liquidity to handle potential redemptions and fair market values should be easily found. Thus, mutual funds that buy into private companies may create a problem for the mutual fund because the investment is not liquid and also because there are issues in valuing these private holdings. Because of the rise in private holdings, Grind [2015] reports the Securities and Exchange Commission (SEC) views liquidity as a prime concern for mutual funds with private valuations right behind.

Since there are issues involved in investing in private companies, the question is why would a mutual fund be interested in doing it? One reason is the shrinking number of public companies to invest in for mutual funds. The *Financial Times* [2018] reports the number of US domestic companies has dropped from 8,090 in 1996 to 4,336 in 2017. However, the US market has grown based on market capitalization, 105 percent of gross domestic product in 1996 but 130 percent of gross domestic product in 2017. The main reasons given for the shrinking number of public companies include more mergers, wealthy mega companies acquiring start-ups firms at early stages, costs associated with public listings, burdensome regulation such as the Sarbanes-Oxley Act of 2002, and the rise of available private funding. The impact of these factors is that those companies that do go public are much larger than companies going public in the past. Bullock and Wigglesworth [2018] indicate that 60 percent of firms that did an IPO raised less than $30 million in the 1980s whereas about 30 percent of all IPOs in the 1990s raised less than $30 million. However, by the 2000s, these small IPOs accounted for less than 10 percent of the total. Clearly small dollar IPOs are simply not an effective means of raising equity capital. In 2017, 189 companies did an IPO and raised nearly $50 billion in total.

The downside of going public later when the firm is larger is that investors are buying into businesses that have already experienced explosive growth in revenues and valuations. If small private companies do not go public, retail investors may well miss the ability to share in the potential performance that is currently only available to a small group of insiders and accredited investors. Bullock and Wigglesworth [2018] quotes Jay Clayton, the chairman of the SEC, at a meeting of Sifma, the main US securities industry group, "It's about improving investor opportunities. A broader portfolio of public companies is important to retail investors. A broader portfolio and [one] more exposed to the growth stage that would be better. If you continue to shrink that pool, you are going to shrink their opportunities." In other words, smaller investors may simply not have access to the private firm growth and hence the surge in valuations.

Foley and Waters [2014] point out there has been tremendous interest in pre-IPO funding due to the performance opportunities. For late stage private companies, asset managers like BlackRock, Fidelity and T Rowe Price have been increasing their investments in private firms trying to get a jump on others who they might have to fight for an allocation at a higher price when the company goes public. Foley and Waters indicate asset managers have been developing systems for assessing private sector investments. The managers are trying to establish better relationships with venture capital firms to gain greater access to even more investment opportunities. They quote Andrew Boyd, head of global equity capital markets at Fidelity, "If you want to be a public company, there is no reason to wait to start talking to the big public company investors. The strongest companies are able to attract long-term investors who will hold on to shares through the IPO, hoping to increase the chances of a successful flotation."

Schwartz [2017] suggests one of the lead mutual funds investing in private companies is Fidelity's flagship Magellan Fund. He points out several problems with mutual funds investing in private companies. These problems include the lack of awareness

among mutual fund investors, lack of liquidity for mutual fund shares, lack of venture-capital expertise among mutual fund management, and the lack of accountability over how funds value their ownership stakes in startups for purposes of calculating their net asset values. All of these things can create an opportunity for management to manipulate such valuation estimates. One conclusion of Schwartz is that the Magellan Fund is manipulating its performance when calculating the net asset value. However, when examining the performance of Magellan's private investments, Schwarz includes the performance of these private investments after the private companies have gone public. We re-examine the performance of the private company holdings by the Magellan Fund through time to see the recorded performance prior to the private holding going public or being sold and the performance of the holdings after the private companies have gone public. We then explore how the Magellan Fund has accounted for the private investments performance and whether the Magellan Fund has truly manipulated the performance of their private holdings in any way to inflate their net asset value. This paper provides a look at just how mutual funds are using private firm investments and why investors in such funds may have a disadvantage in clearly understanding the true valuation.

Rise of the Private Company

Private firms such as Uber Technologies and Airbnb have business models that are disrupting conventional markets and money has poured in. Eule [2016] states the stock market crash of 2008-09 curbed investor enthusiasm for IPOs and new sources of liquidity have emerged for company founders and insiders. In addition, regulatory changes have made it easier to stay private and harder to be public. Eule cites work from Jay Ritter, a University of Florida professor who has studied the IPO market for 35 years. From 1980 to 2000, an average of 310 companies went public every year but since then the average has fallen to 111. Not only has the number of IPOs declined, the number of publicly listed companies has followed the IPO trend as well. Mauboussin, Callahan, and Majd [2017] state the public listings of companies fell around 50 percent in the US from 1996 through 2016, whereas public listings of companies increased about 50 percent in other developed countries. They show that based on GDP, GDP growth, population growth, and measures of corporate governance, the US should have more than 9500 public companies. Among the reasons given for the drop in US IPOs are mergers and acquisitions as well as the increased cost of being a public company due to regulatory requirements. This decline of IPOs has forced mutual funds to look at private firms for growth opportunities. Although it represents only a small percentage of their investments, Fidelity Investments, T. Rowe Price Group and Wellington Management all have sizable dollar stakes in private firms.

In essence, the IPO process has become a strategy that is best avoided for as long as possible. In response to the decline in US IPOs, Congress passed the JOBS Act in 2012 to ease the IPO process and to make it easier for small companies to go public. Eule [2016] states the net effect of the JOBS Act has actually been to stall the market. A key provision of the JOBS Act makes it easier to stay private longer. The JOBS Act allows companies to accumulate up to 2000 private investors (previously it was 500) before being forced to disclose public company type information and the limit no longer includes employees granted stock as compensation. Due to the tech bubble in 2000 and the subsequent accounting related scandals such as Enron, Congress passed the Sarbanes-Oxley Act in 2002, which raised the bar for public company disclosures. Thus, the cost of being public became more expensive. Eule argues that this extra cost provides an incentive for companies to remain private longer as they grow in size.

The institutional investor who wanted exposure to US equities in 1976 would only need to buy a diversified portfolio of public companies and a venture capital (VC) fund. In 2016, investors would need to buy a diversified portfolio of public companies, a private equity fund, and an early-stage as well as a late-stage VC fund. Mauboussin, Callahan, and Majd [2017] indicate that private equity firms manage roughly $825 billion in 2016, up from $80 billion in 1996. The companies publicly listed now are bigger, older, and are in more concentrated sectors than two decades ago. Mauboussin, Callahan, and Majd show the average market cap of a firm in 1996 was $1,683 million and by 2016 the market cap grew to $6,893 million. Thus, individual investors have a limited ability to access the complete US equity market.

Mauboussin, Callahan, and Majd [2017], similar to Eule [2016], discuss reasons why eligible companies do not see a net benefit in publicly listing through an IPO. First, the cost of being a public company has gone up so it only makes sense for larger companies to undertake an IPO. They show the median age of a company doing an IPO has increased from 7.8 years during the 1976 to 1996 period, while the median age increased to 10.7 years from the 1997 to 2016 period. Second, companies today need less human and physical capital. Their examples of this point is comparing Amazon to Wal-Mart and Facebook to Ford. Amazon generated $136 billion of sales in 2016 using $19 billion in invested capital for a capital velocity ratio of 7.1. Wal-Mart had sales of $486 billion but used $135 billion in invested capital for a capital velocity of 3.6 times. Facebook's sales per employee were $1.6 million in 2016 whereas Ford's were $755,000. Third, there is plenty of late-stage funding available. The five largest startup companies with the highest implied valuations have raised over $28 billion. Finally, employees of private companies have the ability to sell shares. Airbnb raised $850 million in the Fall of 2016 and allowed employees to sell $200 million worth of stock. According to *The Wall Street Journal*, as of October 2018, there are 163 venture-backed private companies with a value in excess of $1 billion (see Austin, Canipe and Slobin [2015]).

Companies today are building a lot of value pre-IPO rather than post-IPO. Individual investors that do not have a good access to private capital are potentially missing substantial gains. With value being build up pre-IPO, more mutual funds and hedge funds are seeking to participate in late-stage venture capital funding. Mauboussin, Callahan, and Majd [2017] show that 26 mutual fund families had $11.5 billion invested in late-stage venture companies. Fidelity had 59 funds investing over 45 percent of the $11.5 billion in private companies. Andrew Boyd, head of global equity capital markets at Fidelity, suggested that the pre-IPO market has become the IPO market of the past.

Mutual funds provide daily net asset values so the inclusion of less liquid private firms has led to the net asset values coming under scrutiny. Reichart [2018] reports there are other reasons why private firms have become part of the mutual fund investments. She states private companies have delayed IPOs and by raising the funds privately, managers can build their firms without the pressure of being publicly traded. Second, she indicates some of the private firms have changed the competitive dynamics in their industries and are actually larger than their publicly traded peers. Finally, she argues active managers, who have underperformed against passive funds in the picked over public stock market, like the appeal of the performance edge that private firm investments can potentially offer.

Valuing Private Companies

Since no industry standard exists for valuing private companies' securities, the challenge for mutual funds is trying to come up with a value for the private companies' securities they have invested in for reporting purposes. Reuters [2016] reports valuation committees look at revenue growth, competition, barriers to entry and what others paid in subsequent funding rounds. Reichart [2016] points out there may be a variety of factors that contribute to inconsistent pricing of the private company securities reported in the financial press. Grind [2015] states there are clearly varying levels of access to a private company's management team, board and financial information to help with valuations. The larger mutual funds might have the better connections with the private company's management, thus having an advantage of better information to value the private firm. Reichart [2016] reports that Fidelity (the third biggest fund firm holding US mutual fund assets) may have deeper resources to evaluate private companies. Thus, the biggest challenge mutual funds face is trying to determine the value of their private company investments. Reichart states the biggest fund companies, such as Fidelity, have a committee composed of individuals with backgrounds in accounting, compliance, trading, equity, or fixed-income sets that monitor the valuations of private companies. Anyone at Fidelity that has an interest in seeing investments appreciate such as portfolio managers and analysts cannot sit on the committee to avoid conflicts of interest. Nonetheless, the bottom line is that valuation of the private firm holdings by mutual funds is "soft."

Reichart [2016] indicates the starting point of trying to value a private company's security is to examine the most recent funding round. Funding rounds tend to be infrequent so the fund needs to consider company specific and market based factors as they monitor these private company investments. She points out that stock price movements of publicly traded peers or the market as a whole may trigger a markdown or markup. The problem is there is no industry wide consensus for making valuation changes. Reichart showed an example of the wide range of valuations from Morningstar data that showed Airbnb series D and E shares in the second quarter of 2016 being as low as $88.44 to the high of $130.39. A number of factors contribute to this inconsistent pricing such as the fund companies holding different types of equity. The securities purchased by the mutual funds may be different such as preferred stock, common stock, and possibly convertible bonds that could affect valuations based on payouts, growth potential and priority in default. In addition, the fund companies have varying levels of company specific information so they are making a judgment call based on different time horizons, or they possibly are weighing market and peer group stock price movements differently.

If inconsistent pricing of private firm securities occurs, the fund's net asset values (NAVs) can vary by the amount of the discrepancy. Shareholders transacting at NAV when buying or selling fund shares could end up paying more or less for their proportional share of the same security depending on the particular fund's valuation approach. However, Reichart [2016] points out that the direct impact to a fund owner is likely immaterial given the small role such investments play in virtually all fund portfolios with private firm securities.

Magellan Investments

Mutual funds must publicly disclose their portfolio holdings on a quarterly basis. In their disclosures, mutual funds must identify by name each security owned at period end indicating the number of shares owned or the principal amount for bonds and the value of each security. Mutual funds have to indicate each security that is restricted (not registered with the Securities and Exchange Commission (SEC) under the Securities Act of 1933) and disclose the dollar amount and the percentage of the portfolio invested in restricted securities. By law, mutual funds must pay redemption proceeds to redeeming shareholders within

seven days. Thus, mutual funds historically have not invested much in assets of illiquid securities. SEC guidelines limit mutual funds from investing no more than 15 percent of their assets in restricted securities.

The restricted securities information obtained in this study comes from the quarterly filings of the Magellan Fund on the SEC's Electronic Data Gathering, Analysis and Retrieval (EDGAR) website (https://www.sec.gov/edgar/searchedgar/companysearch.html). The first recorded investment in restricted securities for Fidelity's Magellan Fund was in their annual report, March 31, 2011. Table 1 shows from March 31, 2011 to March 31, 2018, the dollar amount invested in restricted securities grew about 42%, going from 0.09% of the total assets of the portfolio to 1.48% of the total assets of the portfolio. Clearly, this is a small but growing part of the Magellan strategy. However, the total assets of the Magellan Fund actually declined by about 4 percent over the time period from around $23 billion to $17 billion. For each quarter from March 31, 2011 to March 31, 2018, Table 1 reports the quarterly holding period return (HPR) for the restricted securities, the HPR for the Magellan Fund, the HPR for the S&P 500 index and the HPR for Morningstar's Large Growth index (which is the comparison Morningstar performance category for the Magellan Fund). Looking at the mean returns, the restricted securities HPR is larger than the returns for the Magellan Fund, the S&P 500 index and Morningstar's Large Growth index, but the risk for the restricted securities is higher as well. The higher returns for the restricted securities may explain why the Magellan Fund has increased its investment in restricted securities through the years or it may indicate the Magellan Fund is trying to boost their returns with the valuation of restricted security investments.

Reuters [2016] states mutual funds have been boosting their performance with bets on private companies by marking up the returns of the private companies at a far greater rate than the broad stock market. The best quarterly HPR reported for the Magellan Fund was in the June 30, 2017 quarter showing a 48.98% return. Given restricted securities were 1.60% of the Magellan portfolio at this point, 48.98% times 1.60% would mean the restricted securities added 0.78% in relative performance. Given the return of 4.25% for the Magellan Fund in June 30, 2017, this 0.78% helped beat the S&P 500 index return of 2.57% but was below the return of the Morningstar Large Growth index of 5.75%. Overall, there does not appear to be much evidence that the Magellan Fund used the restricted security investments to boost their returns.

Schwartz [2017] reported that the Magellan Fund had been tremendously successful over the time period (approximately May 2, 2012 through March 31, 2016) that he examined for Magellan's private company investments. He calculated the weighted mean annualized return of the different investments to be 42%. The issue with his calculation is the inclusion of some time periods where the private company had already gone public. Table 2 first shows the 17 different private company investments and the 3 other restricted security investments listed under the restricted securities for the Magellan Fund. We calculate the annualized return of the private companies from Magellan's initial investment date through either the date of the IPO, the date Magellan stopped listing the investment in their filings on EDGAR, or March 31, 2018 if Magellan still held the company's security. The first returns shown in Table 2 are under the column, annualized return of firm while private. We find the weighted mean return to be 35.87% which is similar to but lower than Schwartz's findings. In addition, we examined the annualized weighted mean return on the 3 other restricted security investments which are still private, and the return was a negative 29.85%. Adding these 3 investments into the 17 private company returns drops the weighted mean return to 21.54% which is about half the return as reported by Schwartz. Examining the return of the Magellan Fund in the column, annualized return of the Magellan Fund, the mean return of the Magellan Fund over the same time period as the private holdings was 12.68%. Not listed in the table, the S&P 500 annualized returns over the same time period as the private holdings produced a return of 9.94%. Once again, the Magellan Fund investments in restricted securities appear to help the overall return of the fund providing a further reason that restricted securities investments by mutual funds are growing over time.

The best private company investment is the 466.56% return in Mobileye N.V. in August of 2014. In Table 1, the investment in restricted securities in September 30, 2014, was 0.80% of Magellan's total value. Taking 466.56% return times the 0.80% weight in restricted stock, times the weight of the Mobileye N.V. investment of 4.15% indicates the investment only increased the return of Magellan by 0.16%. Once again, this increase does not appear likely to boost the return of the Magellan Fund that the financial press reports.

Next, we looked at the return of each company that went public from its IPO date to either the date Magellan stopped listing the investment in their filings on EDGAR or March 31, 2018, if Magellan still held the company's security. These returns are listed under the column, annualized return of firm after IPO in Table 2. There are 7 private firms that went public and the annualized weighted mean return while Magellan held the securities was 95.97%. This return is higher than the private company weighted return of these 7 companies of 91.12% when the companies were still private. With so few private companies exited so far, it is hard to judge the strategy of the Magellan Fund strategy of buying these restricted securities. However, this does seem to indicate the Magellan Fund is not necessarily boosting their returns with their private company investments but yet the performance is doing well. The investments the Magellan Fund made in these private companies has paid off and explains why the Magellan Fund growth in restricted securities has grown 42% and why Reichart [2016] and Schwartz [2017] indicate larger mutual funds are now increasing their investment in private companies.

Conclusion

Using the Magellan Fund as a case example, we find that the fund has been increasing their investment in restricted securities, mainly in private company securities. We conclude that a major reason the Magellan Fund is increasing its investments in restricted securities is the higher performance of the restricted securities as compared to other investments held in the Magellan Fund. Although there are problems in calculating the value of the private company investments, we do not find any evidence the Magellan Fund is manipulating their private company security returns to report higher overall returns for their fund as some have contended. Again, the valuation of private securities is generally rather soft in terms of true valuation so it is not very clear how Magellan may choose to place a valuation on its holdings of private securities. We might expect though to see other funds, like Magellan, increase their holdings of private securities.

References

Austin, Scott, Chris Canipe and Sarah Slobin. "The Billion Dollar Startup Club." *The Wall Street Journal*. (February 18, 2015), Retrieved from https://www.wsj.com/graphics/billion-dollar-club/.

Bullock, Nicole, and Robin Wigglesworth. "US Stock Markets Seek Depth in IPO Pool." *Financial Times*. (January 8, 2018), Retrieved from https://www.ft.com/content/7457ee7e-f074-11e7-ac08-07c3086a2625.

"At a Record High, the US Market is Still Shrinking." *Financial Times*. (August 24, 2018), Retrieved from https://www.ft.com/content/d743e8ec-a787-11e8-926a-7342fe5e173f.

Chernenko, Sergey, Josh Lerner, and Yao Zeng. "Mutual Funds as Venture Capitalists? Evidence from Unicorns." Working Paper (November 8, 2017), Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2897254.

Eule, Alexander. "The Disappearing IPO Market." *Barron's* 23-25. (June 6, 2016) Retrieved from http://www.barrons.com/articles/are-unicorns-killing-the-2016-ipo-market-1465018470?tesla=y.

Foley, Stephen and Richard Waters. "Early Bird Investors Swoop on Opportunities for Pre-IPO Funding." *Financial Times*. (July 4, 2014), Retrieved from https://www.ft.com/content/6c1f3d20-fe1a-11e3-acf8-00144feab7de.

Gage, Deborah. "The Venture Capital Secret: 3 Out of 4 Start-Ups Fail." *The Wall Street Journal*. (September 20, 2012), Retrieved from http://www.wsj.com/articles/SB10000872396390443720204578004980476429190.

Grind, Kirsten. "Regulators Look Into Mutual Funds' Procedures for Valuing Startups." *The Wall Street Journal*. (November 17, 2015), Retrieved from https://www.wsj.com/articles/regulators-look-into-mutual-funds-procedures-for-valuing-startups-1447796553.

Mauboussin, Michael J., Dan Callahan, and Darius Majd. "The Incredible Shrinking Universe of Stocks, The Causes and Consequences of Fewer U.S. Equities." *Credit Suisse*. (March 22, 2017) Retrieved from https://www.cmgwealth.com/wp-content/uploads/2017/03/document_1072753661.pdf.

Reichart, Katie. "Unicorn Hunting: Large-Cap Funds That Dabble in Private Companies." *Morningstar*. (June 6, 2018), Retrieved from http://corporate1.morningstar.com/ResearchLibrary/article/868704/unicorn-hunting-large-cap-funds-that-dabble-in-private-companies/.

Reichart, Katie. "Unicorn Hunting: Mutual Fund Ownership of Private Companies is a Relevant, but Minor, Concern for Most Investors." *Morningstar*. (December 5, 2016), Retrieved from http://corporate1.morningstar.com/ResearchArticle.aspx?documentId=780716.

Reuters. "These Mutual Funds Have Been Juicing Their Returns With Unicorn Stakes." *Fortune*. (August 11, 2016), Retrieved from http://fortune.com/2016/08/11/mutual-funds-returns-unicorns/.

Schwartz, Jeff. "Should Mutual Funds Invest in Startups? A Case Study of Fidelity Magellan Fund's Investments in Unicorns (and Other Startups) and the Regulatory Implications." *North Carolina Law Review,* Vol. 95, No. 5, (2017 Forthcoming), Retrieved from https://papers.ssrn.com/so13/papers.cfn?abstract_id-2911394.

Appendix

Table 1: Magellan versus Index Returns

Date	$ Amount of Restricted Securities	$ Amount of Restricted Securities / Total Value of Fund	Quarterly HPR on Restricted Securities	Quarterly HPR on Magellan Fund	Quarterly HPR on S&P 500	Quarterly HPR on Morningstar Large Growth Index
3/31/2018	$255,043,000	1.48%	2.43%	1.20%	-1.22%	4.73%
12/31/2017	$248,988,000	1.47%	-2.09%	6.88%	6.12%	6.93%
9/30/2017	$212,837,000	1.30%	1.97%	6.69%	3.96%	5.34%
6/30/2017	$237,663,000	1.60%	48.98%	4.25%	2.57%	5.75%
3/31/2017	$219,988,000	1.43%	4.46%	6.41%	5.53%	8.88%
12/31/2016	$210,595,000	1.42%	-5.32%	2.08%	3.25%	-0.65%
9/30/2016	$222,427,000	1.49%	6.83%	4.62%	3.31%	3.86%
6/30/2016	$214,372,000	1.44%	2.12%	0.69%	1.90%	-0.88%
3/31/2016	$197,967,000	1.30%	-6.00%	-2.14%	0.77%	-1.64%
12/31/2015	$210,600,000	1.33%	-23.09%	8.40%	6.45%	7.25%
9/30/2015	$172,847,000	1.16%	2.38%	-7.61%	-6.94%	-5.53%
6/30/2015	$167,098,000	1.01%	15.12%	1.12%	-0.23%	0.40%
3/31/2015	$146,649,000	0.88%	2.20%	2.83%	0.44%	4.57%
12/31/2014	$133,712,000	0.81%	13.10%	5.22%	4.39%	3.93%
9/30/2014	$130,737,000	0.80%	28.17%	1.62%	0.62%	1.93%
6/30/2014	$143,353,000	0.87%	-1.13%	4.57%	4.69%	7.07%
3/31/2014	$129,816,000	0.81%	36.53%	2.04%	1.30%	-0.37%
12/31/2013	$93,618,000	0.58%	1.52%	10.61%	9.92%	11.47%
9/30/2013	$92,128,000	0.61%	-0.31%	9.05%	4.69%	9.07%
6/30/2013	$81,379,000	0.57%	-18.78%	2.86%	2.36%	0.56%
3/31/2013	$98,678,000	0.67%	0.00%	9.12%	10.03%	6.91%
12/31/2012	$109,038,000	0.77%	1.67%	-0.88%	-1.01%	-3.13%
9/30/2012	$91,082,000	0.60%	0.00%	7.90%	5.76%	6.90%
6/30/2012	$73,745,000	0.50%	0.00%	-5.19%	-3.29%	-4.11%
3/31/2012	$60,179,000	0.37%	0.00%	16.35%	12.00%	17.23%
12/31/2011	$38,753,000	0.25%	0.00%	8.06%	11.15%	9.74%
9/30/2011	$41,465,000	0.27%	0.00%	-19.64%	-14.33%	-11.97%
6/30/2011	$40,512,000	0.20%	0.00%	-3.38%	-0.39%	0.16%
3/31/2011	$19,805,000	0.09%				
Mean	$141,209,448	0.90%	4.21%	2.99%	2.64%	3.37%
Std. Deviation	$71,231,021	0.46%	13.81%	6.75%	5.45%	5.95%

Table 2: Magellan Restricted Security Returns

Company	Initial Investment	Private Company Investments			
		Valuation ($) at IPO Date	Annualized Return of Firm While Private	Annualized Return of Firm After IPO	Annualized Return of the Magellan Fund (same period)
bluebird bio, Inc.	$1,711,000 (July 23, 2012)	$3,078,241 (June 19, 2013)	91.10% (June 19, 2013)	122.66% (Sept. 30, 2015)	24.43%
Cloudflare, Inc.	$3,502,000 (Nov. 5, 2014)		4.08% (March 31, 2018)		11.27%
DocuSign, Inc.	$90,000 (Oct. 21, 2013)		41.58% (March 31, 2018)		12.67%
DocuSign, Inc. Series B	$99,000 (March 3, 2014)		18.23% (March 31, 2018)		11.57%
DocuSign, Inc. Series B-1	$30,000 (March 3, 2014)		18.03% (March 31, 2018)		11.57%
DocuSign, Inc. Series D	$11,000,000 (June 29, 2012)		35.00% (March 31, 2018)		14.54%
	$71,000 (March 3, 2014)		12.66% (March 31, 2018)		14.54%
DocuSign, Inc. Series E	$1,831,000 (March 3, 2014)		18.34% (March 31, 2018)		11.57%
HubSpot, Inc.	$15,000,000 (Oct. 25, 2012)	$22,257,025 (Oct. 9, 2014)	22.35% (Oct. 9, 2014)	46.98% (Dec. 31, 2015)	21.46%
KaloBios (Series E) Pharmaceuticals, Inc.	$8,000,000 (May 2, 2012)	$6,097,168 (Jan. 31, 2013)	-30.36% (Jan. 31, 2013)	-84.69%[a] (Dec. 31, 2013)	10.16%
Malwarebytes Corp. Series B	$35,000,000 (Dec. 21, 2015)		15.70% (March 3, 2018)		14.38%
Meituan Corp. Series D	$10,000,000 (Jan. 26, 2015)		94.90% (Sept. 30, 2015)		-5.81%
Mobileye N.V. Series F	$8,878,000 (Aug. 15, 2013)	$47,061,595 (Aug. 1, 2014)	466.56% (Aug. 1, 2014)	-3.18% (March 31, 2015)	20.68%
Nutanix, Inc. Series E	$6,193,000 (Aug. 26, 2014)	$7,396,528 (Sept. 30, 2016)	8.83% (Sept. 30, 2016)	111.35% (March 31, 2018)	5.17%
Pure Storage, Inc. Series E	$2,121,000 (Aug. 22, 2013)	$5,203,020 (Oct. 7, 2015)	52.51% (October 7, 2015)	-37.50% (March 31, 2016)	12.94%
Roku, Inc. Series F	$11,000,000 (May 7, 2013)	$28,340,278 (Sept. 28, 2017)	24.01% (Sept. 28, 2017)	387.10% (March 31, 2018)	13.95%
Roku, Inc. Series G	$5,000,000 (Oct. 1, 2014)	$8,977,464 (Sept. 28, 2017)	21.59% (Sept. 28, 2017)		10.75%
Uber Technologies, Inc.	$15,000,000 (June 6, 2014)		23.84% (March 31, 2018)		11.52%
Tanium, Inc.	$2,064,000 (April 21, 2017)		25.93% (March 31, 2018)		22.81%
Lyft, Inc.	$40,600,000		0.00%		10.34%

Series H	(Nov. 22, 2017)	(March 31, 2018)		
Ivanplats Ltd.	$11,793,000	-20.46%		13.98%
	(March 28, 2012)	(March 31, 2013)		
WME Entertainment	$15,000,000	13.38%		16.74%
Parent, LLC	(April 13, 2016)	(March 31, 2018)		
China Internet Plus	$10,000,000	19.64%		11.06%
Holdings Ltd.	(Jan. 1, 2015)	(March 31, 2018)		
Weighted Mean		35.87%	95.97%	12.68%
Std. Deviation		97.88%	156.00%	6.22%
		Other Restricted Security Investments		
Rialto Real Estate	$41,200,000	-41.26%		11.62%
Fund LP	(Feb. 24, 2011)	(March 31, 2018)		
RREF CMBS AIV, LP	$17,020,000	3.78%		15.19%
	(Aug. 17, 2012)	(March 31, 2018)		
RREF Midtown	$1,457,000	-100.00%		15.87%
Colony REIT, Inc.	(Dec. 31, 2012)	(March 31, 2018)		
Weighted Mean		-29.85%		12.74%
Std. Deviation		52.04%		2.28%

[a]Fidelity Magellan purchased 366,900 shares at the IPO.

A Simplified Method for State Personal Income Tax Return and Withholding Calculations with S and F Numbers

Robert Kao, School of Business, Park University
John Lee, Rigel Technology Corporation
Richard Lee Kansas State University

Abstract

Taxpayers and state governments prefer an efficient, equity-based, and simpler tax system. In this paper, we assign the tax status and filing period with radical numbers that are used for further simplifying the linear and gradual (LG) state tax system, which has been developed previously for assisting employers to file with precise withholding taxes and submit withholding reports to governments. A great number of taxpayers with non-complex tax situations may have an option to either not file tax returns or file with a simplified process. In addition, governments can verify tax returns with the withholding reports before sending out tax refunds to avoid or reduce tax faults and avoidance. The proposed tax rates have not been greatly varied from the existing rates, which would support state tax administration with an efficient concern and the calculation system can be reduced to a minimum. The research finding reveals simplifications of tax calculations, analyses, reforms, and projections that could reduce the processing time and costs for governments, taxpayers, and businesses significantly.

Introduction and Literature Review

Many U.S. states require two existing personal tax systems. One is the tax withholding schedules and the related withholding tables used for businesses to estimate income withholding taxes. Another is the Tax Table and Tax Computation used for taxpayers to correct the above estimations for filing tax returns. One to five tax statuses are applied in these systems. Both the Withholding Tables and Tax Tables are complex with exceptionally long instructions. Some states still use flat tax rates, which inadequately cover different taxable incomes equitably.

One of the major goals of the 1986 tax reform had been to simplify the tax code or reduce the number of tax brackets. In fact, the 1986 tax law has resulted in more complicated forms and no progress toward a simpler tax code in the later tax bills in 1993, 1997, 2001, 2003, and 2012. Recently, Senator Elizabeth Warren introduced the Tax Filing Simplification Act of 2016 on April 13, 2016. This bill amends the Internal Revenue Code of 1986 to require the Internal Revenue Service (IRS) to establish and operate the following programs free of charge, including online tax preparation and filing software, to provide a program for taxpayers to download third-party provided return information relating to individual income tax returns and to permit individuals with simplified tax situations to elect to have the IRS prepare their returns. However, the IRS may not enter into any agreement that restricts its legal right to provide tax return preparation services, software, or tax return filing services. Individual taxpayers can participate in the programs established by this bill, however must first verify their identity to the IRS. Another tax bill was introduced by House Representative Mike Bishop, which was the Mobile Workforce State Income Tax Simplification Act of 2015 on May 14, 2015, and was passed on September 21, 2016. In this bill, one of the simplification processes is to exempt employers from state income tax withholding and information reporting requirements for employees who are not subject to the income tax.

The complexity of the U.S. tax system has created compliance and equity issues according to the Annual Report to Congress of the National Taxpayer Advocate (TAS Executive Summary 2008). The IRS estimated that individuals and businesses have spent 7.6 billion hours a year or equivalent to 3.8 million full-time workers in one year just to comply with the tax filing requirements. It can convert to $193 billion monetary terms or 14 percent of individual and corporate income tax receipts to fulfill the tax collection laws in 2008. The burden of coping with this complexity falls mainly on the taxpayers and creates high compliance costs. The National Taxpayer Advocate recommended Congress to simplify the tax code extensively. The simplifications include a series of recommendations to repeal the Alternative Minimum Tax, simplify the family status provisions of the Code, streamline education and retirement savings tax incentives, allow taxpayers to exclude modest amounts of canceled debts from income without having to make an affirmative claim, revise the overall penalty structure, and reduce tax sunset and phase out provisions.

Again, the Tax Advocate Service reported to congress (TAS 2012) concerns of the tax complexity issues. The Office of the Taxpayer Advocate in the IRS shows that the U.S. taxpayers, including individuals and businesses, spent more than 6.1 billion hours to complete filings required that contains almost four million words by a tax code. On average, the code has added more than one new provision per day. This resulted in nearly 60 percent of taxpayers hiring tax preparers and another 30 percent relying on commercial software to prepare their returns. The report suggests that the tax laws should be better coherent and

explicit. The computations of tax should be transparent and relatively simple for tax papers' confidence and trust. However, few taxpayers today can confidently understand the full tax code or even that they have correctly computed their tax liabilities. They identify that tax complexity is the most serious problem facing taxpayers and recommend that the Congress urgently simplify the tax code immensely for the ease of tax compliance.

Gardner (2013), in his testimony before the U.S. Senate Committee on Finance, had also recommended the federal government to regulate or restrict state and local governments' ability to raise taxes for coordinating and harmonizing their tax laws to restrict their taxing power and tax systems from becoming more complex. Most states are currently using federal AGI for their own special exemptions, although many of these states still provide fewer exemptions than the federal amounts. Although they provide their own deductions, most of them are linked to federal itemized deductions. Only six states use a slight variation and synchronize their income tax to federal taxable income rather than the AGI. Since taxable income includes federal exemptions and deductions, these states automatically allow the same exemptions and deductions on the federal level. Also, federal exemptions and deductions are indexed with inflation factor. The exemptions and deductions would allow these states to increase with inflation factor and avoid the "hidden tax hikes" each year. If these states have lower exemptions and standard deductions than the federal level, taxpayers are then required to recalculate or take further steps for the difference (ITEP 2011).

Many studies have been devoted to the simplification of personal taxation. Slemrod, et.al. (1994) have investigated two-bracket piecewise linear income tax structures. For promoting the Pareto-efficient tax schedules, they used a two-class economy with at least one marginal tax rate equal to zero and let the marginal tax rate change in either direction. Diamond and Saez (2011) suggested considering the optimal progressivity of earnings taxation and considered whether capital income should be taxed. Freebairn (2012) suggested several personal income taxation reform options. He proposed the removal of tax expenditures for some forms of labor remuneration and the increase of more neutral systems of taxation for different forms of capital income. Davis, et al. (2013) indicated that the use of low income tax credits like the Earned Income Tax Credit (EITC) is an important indicator of tax progressivity. In combination with a flat or only nominally graduated rate structure, they pointed out that these tax breaks can sometimes create an unfair result due to the highest income taxpayers paying less of their income taxes than middle-income taxpayers.

Kao and Lee (2013) proposed a linear and gradual tax system to simplify the existing US progressive personal income rules. The advantage of this new tax structure is to eliminate the current complex Tax Tables and Tax Rate Schedules by using a simpler way to calculate the tax amounts. Kao and Lee (2014) have further developed the LG tax system to simplify current U.S. federal and state corporate income taxation from eight to four federal corporate tax brackets. They suggested that the tax system can also simplify current state individual income systems practically. Several advantages of the LG tax system include the simplified tax/tax rate calculations, analysis, and forecasts with less tax processing time and lower management costs for individuals, corporations, and governments.

This paper applies a new filing status number and filing period number for a further simplification of the LG tax system, which streamlines the LG tax system further to a half page or less for 2015 or current U.S. state personal income taxes. Similarly, a great number of taxpayers with non-complex tax situations may have an option to either not file tax returns or to file with a simplified process. Eventually, governments can verify tax returns with tax withholding reports before sending out tax refunds to avoid or reduce tax evasions and avoidances. In this new proposed method, we will describe the existing state personal tax system, the proposed state tax systems, implications on four selective states, and the benefits from using the new proposed system.

Implications on the State Personal Tax Systems

The Existing State Personal Tax Systems

Each U.S. state has a different tax system. Some states use flat tax rates as their personal tax systems. Some state personal tax systems are complex (such as ones with 6-12 tax brackets). State tax systems for individuals have tax rate ranges (0-14.63%). California personal tax system has its tax rates 1.1%-14.63% with ten brackets, which has the highest tax rate 14.63% in all U.S. states (www.edd.ca.gov/pdf_pub_ctr/15methb.pdf). Iowa personal tax system has nine tax brackets with 0.36%-8.98%. Missouri tax system has tax rates 1.5%-6% with ten tax brackets. Hawaii has tax rates 1.4%-11% with 12 tax brackets.

Each state has different situations. Table 1 shows state tax systems for individuals. Flat tax rates are too simple, which cannot cover all taxable incomes (TI) reasonably. More tax brackets increase complexity of a tax system, which have more smooth tax rate changes. State basic tax brackets and rate ranges are from www.taxadmin.org/assets/docs/Research/Rates/ind_inc.pdf.

We will imply the LG Tax System to four selected states, including California, Hawaii, Missouri, and Kansas for their current tax rate ranges and brackets.

Table 1: Tax Rate Ranges and Tax Brackets in State Tax Systems

State No.	Tax Rate Range	Tax Bracket No.	State	Tax Rate Range	Tax Bracket No.
Alabama	2-5%	3	Alaska	No State Income Tax	
Arizona	2.59-4.54%	5	Arkansas	1-7%	6
Colorado	4.63%	1	Connecticut	3-6.7%	6
Delaware	2.2-6.75%	6	Georgia	1-6%	6
Hawaii	1.4-11%	12	Idaho	1.6-7.4%	7
Missouri	1.4-6%	10	Montana	1.0-6.9%	7
			South Carolina	0-7%	6
			Dist. Of Columbia	4-8.95%	4

The Existing California and the LG Tax System for Its Individual Income Taxes

California's personal income tax system is relatively complex with tax rate range of 1.1%-14.63%, ten tax brackets and five filing statuses as shown in Table 2. The California tax system is more complex than most states. There are two tax systems in the existing California tax system. One tax system is Tax Withholding Schedules, which is used for employers to estimate income tax withholdings. Another tax system is Tax Table and Tax Computation to correct the estimations for tax returns. There are similar tax rates and slightly different taxable income ranges, formulas and taxes in the two tax systems, which are usually changed yearly. California has the most populations out of all the U.S. states.

The 2015 California personal income withholding schedules and withholding tables have 10 pages on weekly, bi-weekly, monthly, quarterly, semi-annually and annually basis (www.edd.ca.gov/pdf_pub_ctr/15methb.pdf and www.ftb.ca.gov/forms/ 2015_California_Tax_Rates_and_Exemptions.shtml#sd). 2015 Tax Table has 5 pages (www.ftb.ca.gov /forms/2015/15_540tt. pdf). Its partial Tax Table is show in Table 2.2. For over 100,000, you must compute your tax using the Tax Rate Schedules.

When the LG tax system is used, the ten tax brackets are simplified to four, with easy taxable income ranges, which are shown in Table 3. The tax rate range check is provided for a self-checking tool as shown in Table 3. The LG tax rate formulas are applied to match the tax rates in Table 2 with simplicity and meaning. For Single or Married/RDP filing separately at TI from 0 to $100,000, a linear relationship of $y = a + x/b$ is found by equation (1) with matched tax rate results.

$$\text{Tax Rate} = 0.01 + TI/1,562,500 \quad \text{(tax rate range: 0.011-0.075)} \tag{1}$$

The slope $1/b = 1/100,000/(0.075-0.011) = 1/1,562,500$ and 0.075 is calculated from (2,463.68+0.1023(100,000-50,869)/ 100,000. For Single or married/RDP filing separately at TI from $100,000 to $250,000, another linear relationship $y = a + bx$ is found. At TI=250,000, (2,463.68+0.1023(100,000-50,869)/100,000=0.91. The existing California tax calculation system with 15 pages can be simplified to less than half a page (Table 3) comparably.

Table 2: The Existing California Tax Withholding Schedules

Taxable income is					
Over	**But not over**	**Tax**			**Amount is over**
Schedule X – Single or Married / RDP filing separately					
$0	$7,749	$0.00	plus	1.10%	$0
$7,749	$18,371	$85.24	plus	2.20%	$7,749
$18,371	$28,371	$318.92	plus	4.40%	$18,371
$28,995	$40,250	$786.38	plus	6.60%	$28,995
$519,687	$1,000,000	$55,568.65	plus	13.53%	$518,687
$1,000,000 and over		$120,555.00	plus	14.63%	$1,000,000
Schedule Y – Married/RDP filing jointly or qualifying widow(er)					
$0	$15,498	$0.00	plus	1.10%	$0
$15,498	$36,742	$149.10	plus	2.20%	$15,498
$1,039,374 and over		$111,570.44	plus	14.63%	$1,039,374
Schedule Y — Head of Household					
$0	$15,508	$0.00	plus	1.10%	$0
$1,000,000 and over		$114,672.71	plus	14.63%	$1,000,000

Table 3: California Personal Tax Table (5 pages)

Least	Over	Status 1 or 3	2 or 5	4
$1	$50	$0	$0	$0
$51	$150	$1	$1	$1
$39,951	$40,050	$1,362	$699	$698
$84,951	$85,050	$5,382	$3,093	$3,944

Filing status: 1 or 3 (Single; Married/RDP Filing Separately), 2 or 5 (Married Filing Jointly; Qualifying Widow(er)), 4 (Head of Household)

Table 4: The LG Tax System for California Individual Tax Rates with Range Check

TI×F/S	Yearly Taxable Income (TI)/S	TI	LG Tax Rate (TR) Formula TR	TR Check	Tax: TI × Rate
	0-100,000		0.011+TIxF/1,562,500/S	0.011-0.075	
	100,000-250,000		0.06433+TIxF/9,375,000/S	0.075-0.091	
	250,000-500,000		0.0758+TIxF/16,447,368/S	0.091-0.1062	
	500,000		0.1463 – 20,050xS/TI/F	0.1062-0.1463	

For taxpayers, a status number (S) is 1 for Single (or Married filing separately), 2 for Married filing jointly (or qualifying widow(er)) or =1.5 for Head of household. Filing period (F) is 1 (yearly Tax Return), 2 (semi-yearly), 4 (quarterly), 12 (monthly), 24 (semi-monthly), 26 (bi-weekly) or 52 (weekly) or 365 (daily). For employers to file withholding, S and F numbers are used. Computer programs for withholding taxes, tax withholding reports and tax returns have been designed to do these automatically.

There are only 1.6% (0.091-0.075) and 1.52% (0.1062-0.091) between $100,000-250,000-500,000. The 4 tax brackets can be further simplified to 3 reasonably by Table 4. The existing CA tax system and the LG tax system are comparable, which are shown in Figure 1. When tax information such as total incomes, filing status, deduction and exemption are inputted, taxable income and tax can be calculated automatically. For Single or married filing separately, total tax can be calculated from the equation (2). Here w, x and y are individual numbers during the four taxable income ranges. Other similar equation(s) may be used for tax analysis and projection.

$$\text{Total Tax} = 0.011\sum TIw + \sum TIw^2/1,562,500 + 0.0758\sum TIx + \sum TIx^2/12,820,513 + 0.1463\sum TIy-20,050y \qquad (2)$$

Table 5: The LG Tax System for California Individuals with S and F numbers (3 tax brackets)

TI×F/S	Taxable Income TI×F/S	TI	LG Tax Rate (TR) Formula TR	TR Check	Tax: TI × Rate
	0-100,000		0.011+TI×F/1,562,500/S	0.011-0.075	
	100,000-500,000		0.0672+TI×F/12,820,513/S	0.075-0.1062	
	500,000		0.1463 – 20,050×S/TI/F	0.1062-0.1463	

Figure 1: Comparison of Tax Rates between the Existing and the LG Tax System (10/3 tax brackets)
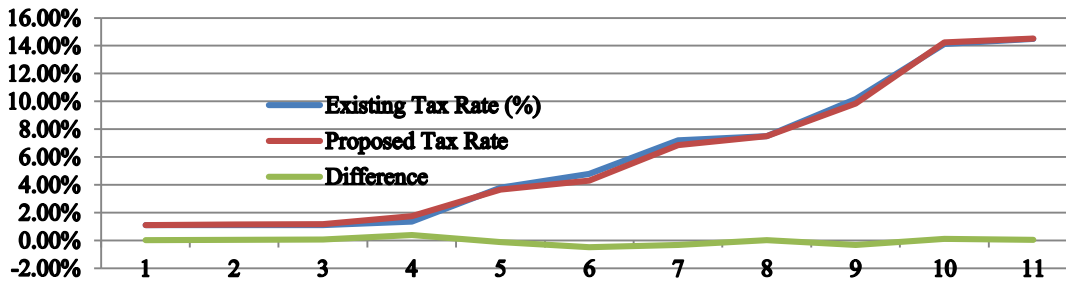
Taxable income: 1=$100, 2=$500, 3=$1,000, 4=$10,000, 5=$40,000, 6=$50,000, 7=$90,000, 8=$100,000, 9=$500,000, 10=$5,000,000, 11=$20,000,000)

The LG tax rate formulas in Table 4 or 5 are much simpler than the existing withholding and tax tables, which reduces from 15 pages to ~1/4 of a page. The tax processing time and costs are then reduced. When employers, taxpayers, and the government use the same tax system (Table 3 or 4) for withholding taxes, tax returns and tax analysis, accurate withholding taxes may be reached for many taxpayers with non-complex tax situations such as one-source incomes and standard deductions. Then these

taxpayers may not need to file normal tax returns because of accurate withholding taxes by accurate information adjusted by the end of a year. Tax withholding reports, which cover detailed tax data such as name, social security number, income, deduction, retirement, credit, taxable income, tax rate and tax, from employers may be done from payroll summaries by January 31with or without modification. For a tax reform, tax brackets are 3 and taxable income ranges are fixed. Lawmakers adjust tax rate top/bottom ranges to meet a tax goal. A government can use tax withholding reports compared with tax returns before sending out tax refunds, which can reduce tax filing mistakes and fraud crimes.

Hawaii Tax System Simplification with the LG Tax System

The 2015 HI personal tax system has the most tax brackets (12) out of all the U.S. states. Its tax rates are from 1.4% to 11% with Tax Table (12 pages) and the three filling statuses in Table 6 (http://files.hawaii.gov/tax/forms/2015/n11ins.pdf), which is the most complex state tax system. When the LG Tax System is used, the existing system can be simplified significantly with 4, 3 or 2 tax brackets and without Withholding Table and Tax Table. The LG tax rate formulas with tax rate range checks are used to replace the above Tax Table and Tax Rate Schedules with less than half of a page. Figure 2 shows their tax rate differences from the two tax systems, which are very comparable to each other. The LG tax system is much easier to use than the existing tax system.

Table 6: Existing HI Tax Rate Schedules (12 tax brackets)

I. Single and Married Filing Separately		II. Married Filing Jointly		III. Unmarried Head of Household	
0 - $2,400	1.4% of TI	0 - 4,800	1.4% of TI	0 - $3,600	1.4% of TI
$2,400 - $4,800	34+3.2% (TI-2400)	4,800-9,600	67+3.2% (TI-4800)	3,600-7,200	50+3.2%(TI-3600)
175,000-200,000	13,879+10% (TI-175000)			363,500-300,000	20,818+10% (TI-262,500)
Over 200,000	16379+11% (TI-200,000)	Over 400,000	Over 300,000		24568+11% (TI-300,000)

When S and F numbers are used, the existing complex HI tax system can be simplified from 12 tax brackets to 4 (Table 7), which can be used by all parties. Status number (S) is 1 for Single or Married filing separately, 2 for Y-Married filing jointly (or qualifying widow(er)) or =1.5 for Head of household. Filing period (F) is 1, 2, 4, 12, 24, 26, 52 or 365.

Table 7: The LG Tax System for HI Individuals with S and F factors (4 tax brackets)

TI×F/S	Taxable Income TI×F/S	Your TI	LG Tax Rate Formula	Tax Rate	Tax Rate Check	Tax Rate × TI
	0-20,000		0.014+TI×F/512,820.5/S		0.014-0.053	
	20,000-50,000		0.04375+TI×F/2,097,902/S		0.014-0.053	
	50,000-200,000		0.06281+TI×F/10,478,519/S		0.014-0.053	
	over 200,000		0.11-5,600×S/TI/F		0.014-0.053	

Figure 2: Tax Rate Differences between the Existing Hawaii and the LG Tax System
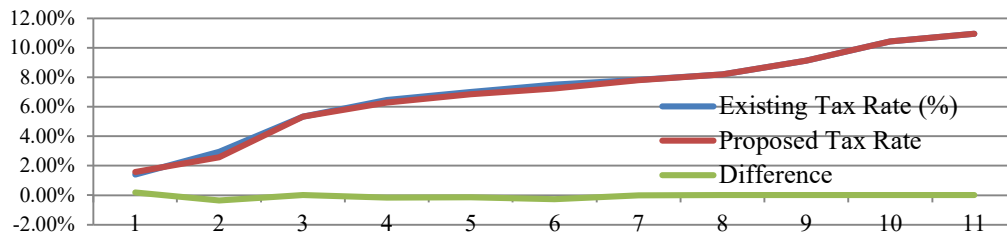
Taxable income: 1=$1,000, 2=$6,000, 3=$20,000, 4=$40,000, 5=$60,000, 6=$100,000, 7=$160,000, 8=$200,000, 9=$300,000, 10=$1,000,000, 11=$20,000,000

There are only 1.5% (0.068-0.053) and 1.4% (0.082-0.068) between $20,000-50,000-200,000. When the two tax brackets are simplified and combined into one range with 2.9% tax rate difference, the four tax brackets can be further simplified to three reasonably and practically. Their tax rate formula becomes 0.04978+TI×F/6,206,896.6/S (0.053-0.082) for $20,000-200,000. The changes are shown in Table 7.

Table 8: The LG Tax System for HI Individuals with S and F factors (3 tax brackets)

TI×F/S	Taxable Income TI×F/S	Your TI	LG Tax Rate Formula	Tax Rate	Tax Rate Check	Tax Rate × TI
	0-20,000		0.014+TI×F/512,820.5×S		0.014-0.053	
	20,000-200,000		0.04978+TI×F/6,206,896.6×S		0.0053-0.082	
	over 200,000		0.11-5,600×S/TI×F		0.082-0.11	

For the tax analysis, projection and reform of the filing status I, total tax equation is as following equation (3), which is used to calculate the total tax for the whole two groups of Single and Married taxpayers filing separately. The total tax equation or similar equations may be used to do total tax, tax difference, average tax or tax rate and tax data analysis and tax projection simply and practically, which do not need to have individual tax data at first and then to add them together. Tax rate or tax is a simple function of TI.

$$\text{Total Tax} = 0.014\sum TIe + \sum TIe^2/512,820.5 + 0.04978\sum TIf + \sum TIf^2/6,206,896.6 + 0.11\sum TIg-5600g \qquad (3)$$

The Existing Missouri and the LG Tax System for Individual Taxes

The Missouri personal income tax system has 10 tax brackets with marginal tax rates at 1.5%, 2%, 2.5%, 3%, 3.5%, 4%, 4.5%, 5%, 5.5% and 6% and Tax Table for individuals. All individuals are required to search and calculate their total income taxes. Tax numbers in the Tax Table are from one number to another number without smooth changes gradually. Tax withholding tables for employers to use have 10 pages. The 10 tax brackets and 10-page tax withholding tables are complex. When the LG tax system and F and S numbers are used, the two tax rate formulas are found and matched (Table 9), which are simple and reasonable for all related parties to use the same tax system (Table 9). S number is 1 for different MO tax statuses. Then tax return, analysis, reform and projection can be done simply and practically. Tax rates change smoothly and continuously. There is almost no tax rate difference by LG and existing tax systems compatibly.

Table 9: The LG Tax System for MO Individuals with S and F Factors (2 tax brackets)

TI×F/S	Taxable Income TI×F/S	Your TI	LG Tax Rate Formula	Tax Rate	Tax Rate Check	Tax Rate × TI
	0-9,000		0.015+TI×F/450,000		0.015-0.035	
	over 90,000		0.06 – 225/TI×F		0.035-0.06	

Tax calculation, analysis, reform and projection become easy to do. Total tax is calculated with the equation (4). Here m and n are single individual numbers during the two taxable income ranges.

$$\text{Total Tax} = 0.015\sum TIm + \sum (TI)^2m/450,000 + 0.06\sum TIn-225n \qquad (4)$$

The LG Tax System for Kansas Individual Tax Returns

Kansas has relatively simple tax system with 3 tax brackets. But its existing tax withholding tables have 22 pages and Tax Table (for TI ≤$100,000) has 8 pages, which are complex. Married filing separately, Single, and Head of Household are in the same tax status with S number 1. Table 10 is simple tax return form (half page) for all taxpayers (F=1) with the LG tax system. For employers, different filing period (F) is 1 (yearly Tax Return), 2 (semi-yearly), 4 (quarterly), 12 (monthly), 24 (semi-monthly), 26 (bi-weekly), 52 (weekly) or 365 (daily). Lines 6-7 in Table 9 are used to figure out income withholding taxes. It has been reported from KS Department of Revenue that about 80% of Kansans use the standard deduction now.

Tax calculation, tax return, analysis, reform and projection become easy to do. Total tax is calculated with such as the equation (5). Here p and q are married filing separately numbers during the two TI ranges.

$$\text{Total Tax} = 0.031\sum Tip + \sum (TI)^2p/2,967,359 + 0.057\sum TIq – 457.5q \qquad (5)$$

Final income tax (C9) = Income Tax-KS state EIC-Nonrefundable tax credits (If Final income tax is less than 0, enter 0).

Tax (+Owe/-Refund) (F9/G9)=Final income tax-Refundable tax credits–KS tax withheld–Tax refund (last year) (6)

G9 is negative for tax refund (providing bank information) or positive for owned tax (asking tax payment). If your tax refund (G9) is less than $100, it may be delay to next year (into F4) to save tax processing time and costs. If your taxable income is less than $100,000 with standard deductions, you may use your tax withholding report (similar to W-2 with more details) with or without modification and attach necessary documents to replace normal tax return.

Table 10: 20XX KS Tax Return Form by the LG Tax System

	Married filing separately	Married filing jointly or qualifying widow(er)	Single	Head of household	
Tax status number	1	2	1	1	Record's
Standard reduction	3,750	7,500	3,000	5,500	Barcode

Taxable income (TI) = State adjustable gross income/Modification-Deduction-Exemption

Exemption: 2,250/person. Additional exemption for blind: 850/person or for 65 or older: 850/person.

A	B	C	D	E	F	G	
Year	Your Name	SS #	Spouse Name	Spouse SS #		Child #	1
2020							2
Status (S)	Federal AGI	Modifications	Standard/Itemized KS deductions	KS exemptions	Tax refund (≤100/last year)	Taxable income	3
							4
TI/S	Yearly taxable income (TI)/S	Your TI	LG tax rate formula	Tax rate check range	Tax rate	Income Tax	5
	0-50,000		$0.031+TI/2,967,359/S$	0.031-0.04785			6
	50,000		$0.057-457.5 \times S/TI$	0.04785-0.057			7
KS EIC	Nonrefundable tax credits	Final income tax (≥0)	Refundable tax credits	KS tax withheld	Tax (Owe)	Tax (Refund)	8
							9

The LG tax system with 2 or 3 (preferred 2) tax brackets can replace the existing state tax systems with different tax brackets (up to 12) and simplify existing tax complexity with smooth tax rates. A flat tax rate is easy but too simple, which cannot cover all taxable incomes reasonably. More tax rate brackets (4-12) can increase complexity of a tax system, which need more processing time and costs. The tax status and filing period numbers are used to simplify existing tax systems, which can be applied for many U.S. states. Also the LG tax system with the tax status and filing period numbers can be applied for the U.S. federal tax system and other countries.

For a tax reform, lawmakers usually struggle to select the number of tax brackets (1-12), taxable income ranges, tax rates, and computations at the very least to meet a tax goal. It is difficult to consider several factors with different options at the same time. With the LG tax system, which provides the most reasonable, simple and fair tax rates, these factors and options are fixed and only tax rate ranges (bottom/top) are adjusted to meet a tax goal. Tax rate formulas are decided by fixed taxable income and tax rate ranges. Then "complex" political factors and options are converted into simple technical issues.

Conclusion

The proposed method in this paper could simplify and combine the tax status number, taxable income, tax rate formula, filing period number, tax rate check, and tax rate into a short tax simplification table. Tax status and filing period number can be used to further simplify the LG state tax system and reduce to two or three taxable income ranges. The LG state tax system can help employers to file the withholding taxes accurately for taxpayers and provide governments tax withholding reports earlier by January 31. Many taxpayers with non-complex tax situations can exempt to file their normal tax returns with or without modifying their tax withholding reports. Governments could verify tax returns with tax withholding reports before sending out tax refunds to reduce tax faults and possible fraud crimes. The related software programs for calculating the withholding tax return can be designed to process it automatically. The proposed tax rates are compatible with the existing tax rates, which can be complied with the current tax administration.

The LG tax system simplifies and combines existing state personal tax systems with tax withholding schedules/related withholding tables for businesses and Tax Table/Tax Computation for taxpayers into one simple tax system for many states. All related parties can benefit from using the same simple LG tax system, and the tax rates could be easily checked by tax officers and taxpayers. This can also be complied with the enacted Senate Bill (2789) of Tax Filing Simplification Act of 2016. The proposed tax system would provide more reasonable and fair tax rates that could allow lawmakers to help tax reforms by

adjusting tax rate ranges and other related factors to meet the tax goal. The potential contributions to this research finding could simplify tax calculation, tax return, analysis, reform, and projection. The tax processing time and costs for lawmakers, governments, taxpayers, and businesses can be reduced significantly.

References

Bruce, Donald, William F. Fox, and Zhou Yang. 2010. "Base Mobility and State Personal Income Taxes," *National Tax Journal*, 63, (4, Part 2), 945–966, December.

Davis, Carl, Kelly Davis, Matthew Gardner, Robert S. McIntyre, Jeff McLynch, Alla Sapozhnikova (2013). "Who Pays? A Distributional Analysis of the Tax Systems in All 50 States," 4th Edition, the Institute on Taxation and Economic Policy (ITEP), January.

Diamond, Peter and Emmanuel Saez. 2011. "The Case for a Progressive Tax: From Basic Research to Policy Recommendations," CESIFO Working Paper NO. 3548, Category 1: Public Finance, August.

Economic Growth and Tax Relief Reconciliation Act of 2001 (EGTRRA 2001), Public Law 107–16, 115 Stat. 38, June 7.

Freebairn, John. 2012. "Personal Income Taxation," The Economic Society of Australia, Economic Papers, Volume. 30, NO. 1, 18–23, March.

Gardner, Matthew. 2013. "Tax Reform: What It Means for State and Local Tax and Fiscal Policy," the Institute on Taxation and Economic Policy, Testimony before the Senate Committee on Finance, United States Senate for Hearing, March 19.

Hyman, David N. 2011. "Public Finance, A Contemporary Application of Theory to Policy," 10th Ed., South-Western, Cengage Learning.

Institution on Taxation and Economic Policy. 2011. "Income Tax Simplification: How to Achieve It," Policy Brief, July.

Kao, Robert and John Lee. 2013. "The U.S. Personal Income Tax Reform: Linear and Gradual Tax Simplifications," *Academy of Economics and Finance Journal*, Volume 4, 47-55.

Kao, Robert and John Lee. 2014. "The U.S. State Personal Tax Simplifications with the LG Tax System" *Academy of Economics and Finance Journal*, Volume 5, 59-67.

Mobile Workforce State Income Tax Simplification Act of 2015 (H.R.2315). 2016. 114th Congress (2015-2016), introduced May 14, 2015 and passed September 22.

Slemrod, Joel, Shlomo Yitzhak,i and Joram Mayshar. 1994. "The optimal two-bracket linear income tax," Journal of Public Economics 53: 269-290. North-Holland, The Prime Minister's Office, Stockholm S-103 33, Sweden, August.

Tax Guide 2011 for Individuals, Internal Revenue Service, Department of Treasury, Federal Catalog Number 10311G (http://www.irs.gov/pub/irs-pdf/p17.pdf).

Taxpayer Relief Act of 1997, Pub. L. 105–34, H.R. 2014, 111 Stat. 787, enacted August 5.

The American Recovery and Reinvestment Act of 2009 (ARRA 2009), Public Law 111–5, February 17.

The Jobs and Growth Tax Relief Reconciliation Act of 2003 (JGTRRA 2003), Public Law 108–27, 117 Stat. 752, May 23.

The Taxpayer Relief Act of 1997 (TRA 1997), Public Law 105–34, H.R. 2014, 111 Stat. 787, August 5.

Tax Brackets and Tax Rates 2012, U.S. Tax Center, Elizabeth Rosen, http://www.irs.com/articles/tax-brackets-and-tax-rates, January 23.

Tax Filing Simplification Act of 2016 (S.2789), 114th Congress (2015-2016), enacted April, 13.

The National Taxpayer Advocate, 2008. Annual Report to Congress, Executive Summary, December 31.

The National Taxpayer Advocate, 2012. Annual Report to Congress, Executive Summary: Preface & Highlights, December 31.

Six Years After The Storm: Superstorm Sandy's Lingering Impact

Richard Vogel, Farmingdale State College

Abstract

Superstorm Sandy struck Long Island and the New York metropolitan area in October 2012, causing extensive damage to the region, including large scale power outages and significant damages to parts of the commuter rail and subway systems. This paper reviews the region's recovery from the event and provides a preliminary evaluation of the storm's lingering impacts on employment and Long Island's economy.

Introduction

Superstorm Sandy struck the New York metropolitan area on October 29, 2012, late in the hurricane season. The storm caused over $70 billion in damages to the region, causing significant impact to low-lying coastal communities and the region's infrastructure. Parts of Long Island were left without electric power for several weeks and the reconstruction of housing and transportation infrastructure in some parts of Long Island and New York City are still ongoing. Long awaited repairs from damage caused by Sandy to some tunnels used by the NY subway system are only now about to get underway. This paper evaluates the lingering impact of Superstorm Sandy on New York's suburban community of Long Island (Nassau and Suffolk County).

Ortega and Taspinar's (2018) recent study of post-Sandy housing markets in NYC found that there is a continuing price penalty on homes in flood zones of approximately 9.4 percent. Their study rooted in a difference in difference analysis of housing sales data and combined with FEMA damage estimates concludes that this price penalty was the result of factors such as the long lag time involved in the reconstruction of some impacted neighborhoods, changes in perceptions of the risks associated with living in flood zones, and more importantly, changes in the flood insurance programs. Housing neighborhood recovery has been an ongoing issue in the post-Sandy era, as there are still some parts of Long Island where homeowners have been unable to fully complete repairs to their homes and communities are still dealing with post-disaster related problems.

Studies such as Grube, Fike and Storr (2017) have found differential levels of disaster assistance funding provided to individuals and households through FEMA's Individual and Household Program (IHP) varied significantly based upon not only the actual level of impact and damage households suffered, but also based upon factors such as educational level of individuals applying for support and whether one was foreign born. Aid has varied significantly across the various communities on LI based upon factors such as income and insurance coverage, with more affluent households better able to secure reconstruction funding (Greenburg 2014; Faber 2018). Neuman et al (2015) found that there are significant costs associated with both storm surge and climate change induced sea-level rise that suggest that coastal regions as the NY metropolitan area will need to adapt to the changing environment over the next century.

Other studies such as Park et al (2017) found that Superstorm Sandy caused $2.8 billion in damages to lost production (interindustry and interregional) and trade in the first four days of the disaster. Yang and Jahan's (2018) study focused upon New Jersey coastal tourist communities found that there are significant issues associated with full recovery which includes not only repairs and replacement of housing but of tourism infrastructure.

The rest of this paper is organized as follows. In Section 2, Sandy's overall impact on LI is reviewed and documented. Section 3 provides an overview of the region's economy and presents various data associated with the economy. In Section 4, an analysis of the data is presented. The conclusions of the study are presented in Section 5.

Review of Sandy's Impact

Aerts et al (2013) estimated the total impacts of Sandy on NYC at approximately $28 billion. This figure includes both direct and indirect costs, as with any disaster a full accounting of costs must account for both physical impacts and a range of indirect and induced impacts that occur as a result of household and business disruption. On Long Island, just to the east of the city, approximately 100,000 homes and residences were damaged by the storm's impact and 2000 homes were destroyed (Blake et al 2013). Widespread power outages across the Island that lasted for several weeks in some communities, left many households without heat and caused severe fuel shortages as very few service stations had backup generators (FEMA 2017). In nearby Connecticut, power outages lasted for as long as 9 days following Sandy (Wanik et al 2018).

The level of impact from Sandy lead the governor of NY to create the New York Rising Community Reconstruction Program to coordinate, oversee and facilitate the post-Sandy reconstruction process. This office coordinated state assistance and funding processes to households and communities in order to facilitate reconstruction and to create more resilient communities.

As of October 2017, Roy (2017) reported that of the 11,000 homes included under the New York Rising program on Long Island, 3400 of them had still not been fully repaired. There was still $1.4 in repairs and upgrades that needed to be completed to Long Island's power grid to better enable the region to withstand future events. In the immediate aftermath of Sandy, some LI industries such as the recreational fishing industry suffered losses in excess of 50 percent of their normal operating revenues and employed less than 50 percent of their normal levels, directly attributable to post-Sandy business interruption. However, most of Sandy's impacts on LI were on housing and the residential sector (including beach front communities) and coastal community infrastructure. There was some significant damage to the Bay Park sewage facility, the major treatment facility for LI, and it did take significant time to repair and strengthen the facility against future events (an 18 foot barrier was built around it).

Given that most damages and impact were to the residential sector, it is not clear how much LI's economy has changed as a result of Sandy. Institutionally, state and local government has responded by strengthening zoning and building codes to better account for the disaster threat. Flood insurance and hazard maps have been revised to reflect the disaster threat as well. This has resulted in requirements for some homes and properties being forced to add elevation before being allowed to rebuild and added costs to some coastal property owners. LI's economy in the wake of superstorm Sandy is the topic of Section 3.

Overview of the Region's Economy

The New York Federal Reserve's (2018) profile of LI shows a region with a population of just under 2.9 million people, with median household income of $98,000 annually, and the median price of a home at $400,000. Over the past ten years, LI's population has grown by 2.3 percent, and the number of jobs has grown by 6.3 percent. The percentage of the population with a college degree (Bachelor's level) stood at 39.7 percent. While population and job growth are significantly lower than New York State's at 3.7 percent and 8.3 percent respectively, median home price, household income and education level are significantly higher than both state and national levels.

Table 1: Demographic and Social Profile of Long Island

	2009	2010	2011	2012	2013	2014	2015	2016
Population (ACS)	2,865,169	2,811,631	2,820,124	2,831,072	2,839,568	2,850,974	2,855,985	2,854,931
Male	1,403,994	1,371,422	1,376,175	1,381,898	1,386,771	1,392,842	1,395,948	1,395,512
Female	1,461,175	1,440,209	1,443,949	1,449,174	1,452,797	1,458,132	1,460,037	1,459,419
Median Age	39.8	39.9	40.2	40.5	40.6	40.8	40.9	41.1
White	2,292,668	2,214,444	2,204,654	2,201,281	2,185,220	2,173,459	2,166,688	2,157,972
Black/African American	256,029	253,190	255,801	259,249	261,638	265,159	269,360	269,875
Asian	142,290	149,615	152,619	155,774	159,937	165,020	172,008	175,033
Hispanic or Latino	363,666	409,470	425,797	441,097	454,519	469,173	480,768	490,484
Labor Force	1,473,161	1,463,450	1,473,679	1,484,603	1,489,338	1,493,164	1,497,597	1,498,361
Veterans, Age 18-64	80,944	73,744	68,832	64,758	59,150	52,627	46,465	41,248
Median Household Income	$88,303	$88,805	$91,260	$92,150	$92,434	$93,067	$93,199	$95,220
Per Capita Income	$38,121	$38,897	$39,779	$39,404	$39,526	$40,043	$40,277	$41,521
Poverty Level	142,957	148,494	152,462	165,771	173,729	184,047	185,415	186,781
No H.S. Diploma	124,179	126,453	125,993	125,527	126,010	128,535	126,157	126,298
High School Graduate	396,030	381,842	376,638	372,814	371,712	367,065	363,907	355,691
Some College	265,916	262,276	265,136	268,130	264,529	260,757	260,716	258,894
Associate's Degree	145,501	143,055	143,528	144,775	147,574	145,387	143,486	142,614
Bachelor's Degree	338,393	335,853	340,483	341,616	343,056	349,117	354,125	355,989
Postgraduate Degree	254,207	254,578	256,197	258,715	260,183	265,179	266,775	269,058
Total Housing Units	1,002,254	1,031,405	1,035,263	1,035,988	1,036,437	1,037,743	1,037,450	1,037,591
Median House Value	$458,773	$454,399	$443,089	$429,559	$416,978	$410,383	$406,900	$410,500
Foreign Born	465,113	482,668	489,989	499,675	503,299	514,379	520,685	526,899

Source: Jobs EQ

While the figures differ slightly from the Federal Reserve's profile, data from the Jobs EQ database provides a greater breakdown of LI's demographic and social profile. As shown in Table 1, over the past decade, the median age on LI has increased and the population is becoming much more diverse. Over an eight-year period the number of housing units increased

by approximately 3.5 percent, from 1.002 million to 1.037 million. The data in Table 1 also does show that median home prices fell significantly over the eight-year period, partly as a result of the Great Recession, but may also reflect the legacy of Superstorm Sandy (see Ortega and Taspinar 2018). Poverty levels on LI have increased significantly across the period, increasing by 30 percent between 2008 and 2016, from just under 5 percent of the population to 6.5 percent of LI's 2.854 million people.

In terms of structure, LI's economy has changed dramatically over the past thirty years. As shown in Table 2, from January 1990 to December 2018, as a percentage of total employment manufacturing employment has decreased from over 13 percent of total employment to 5.1 percent. Employment in the education and healthcare sector increased from 12 percent of total employment to just over 20 percent of total employment. In the trade sectors, employment fell from 6.3 percent to 5 percent of total employment in wholesale trade and 13.4 percent to 12.25 percent in retail trade. Employment in the information sector fell from 3 percent of total employment to 1.2 percent. Leisure and hospitality employment grew from 6.3 percent to 9 percent of total employment. Government employment fell from 16.6 percent to 14.5 percent of total employment. Professional and business services increased from 10 percent to 12.9 percent of total employment. Mining, logging and construction employment (primarily represented by construction) grew from 4.5 percent to 6.5 percent of total employment. Other services employment increased from 3.4 percent to 4.4 percent of total employment.

Table 2: Distribution of sectoral employment, 1990 to 2018

	1/1990	1/1994	1/1998	1/2002	1/2006	1/2010	1/2012	1/2013	1/2014	1/2018	12/2018
Construction	4.59%	3.26%	4.03%	4.90%	5.19%	4.69%	4.68%	5.02%	5.04%	5.69%	6.51%
Manufacturing	13.00%	9.96%	9.31%	7.85%	7.07%	6.08%	5.97%	5.89%	5.77%	5.37%	5.13%
Trade	23.03%	23.38%	23.13%	22.28%	22.11%	21.20%	21.34%	21.14%	21.46%	21.28%	20.80%
Wholesale	6.39%	6.13%	6.05%	6.00%	5.89%	5.57%	5.51%	5.50%	5.53%	5.35%	5.07%
Retail Trade	13.45%	13.89%	13.81%	13.26%	13.27%	12.65%	12.77%	12.54%	12.68%	12.50%	12.25%
Information	3.01%	2.64%	2.67%	2.78%	2.41%	2.17%	1.96%	1.95%	1.80%	1.37%	1.28%
Professional and Business Services	10.03%	11.36%	12.16%	12.50%	12.30%	12.08%	12.60%	12.86%	12.84%	12.87%	12.99%
Education and Health Services	12.06%	14.65%	15.40%	15.59%	16.66%	18.73%	19.06%	19.07%	19.20%	19.84%	20.24%
Leisure & Hospitality	6.35%	6.38%	6.31%	6.68%	7.11%	7.44%	7.89%	8.17%	8.47%	8.87%	9.07%
Other Services	3.46%	3.67%	3.89%	4.10%	4.20%	4.32%	4.38%	4.37%	4.43%	4.49%	4.44%
Government	16.61%	16.85%	15.94%	16.51%	16.38%	17.46%	16.31%	15.77%	15.24%	14.83%	14.51%

Source: Bureau of Labor Statistics, State and Area Employment, Hours, and Earnings

LI's economy has shifted over the past thirty years reflecting the overall US economy as well. Through the late 1980s, LI hosted an extensive array of defense related manufacturing and service related facilities. The end of the Cold War and the allure of lower cost production opportunities in other parts of the country (and globalization) the region's manufacturing base declined. With the continued rise of a more technology-based economy and the realignment of production on a more globalized scale, the region's economy has continued change.

Analysis

Using data from the Bureau of Economic Analysis, Bureau of Labor Statistics, and Federal Reserve Bank of St. Louis (FRED), we analyze the long-term impact of Superstorm Sandy on employment on Long Island. The analysis is conducted using a simple time series analysis taking into account a time trend and instrumenting for the hurricane. This analysis is preliminary and is presented for the major sectors of the region. The data was found to be cointegrated, and equations presented are evaluated using fully modified least squares (FMLS) and estimated using Eviews 10.

The variables (Table 3) used in the analysis include total non-farm employment, and employment for the following sectors: construction, manufacturing, wholesale trade, retail trade, leisure and hospitality, education and healthcare, professional and business services, information, other services, and government. In addition, estimates of monthly real gross domestic product estimates from IHS Markets was utilized in the analysis. Hurricane and tropical storm data was collected from the National Hurricane Center and converted into two dummy variables, one for the frequency of hurricanes, tropical, and subtropical storms to affect or impact LI from 1990 to 2018, and the other to capture Superstorm Sandy and is zero until September 2012 and 1 from October 2012 forward. In addition, New York's minimum wage was collected and included in the estimations.

The following equation was estimated for each sector in both level and percentage forms: $Employment_{ij} = f(USRGDP_i, NYMinWage_i, Sandy, LITropHur_i, Trend, C)$, where the subscript "i" represents time, and "j" represents sector. Preliminary evaluation of the data did indicate that there are some stationarity issues with some of the data, so many of the inferences and conclusions from the analysis are still tentative.

Table 3: Variables used in the analysis

Variable name	Description	Time period	Source
USRGDP	U.S. real GDP	1/1992-10/2018	IHS Market
LITROPHUR	Dummy variable for tropical storms and hurricanes impacting Long Island	1/1990-12/2018	National Hurricane Center
Sandy	Dummy variable indicating impact of Superstorm Sandy, "0" before 10/2012 and "1" from 10/2012 forward	1/1990-12/2018	
NYMinWage	New York State minimum wage rate	1/1990-12/2018	New York State Department of Labor
Construction	Construction employment	1/1990-12/2018	BLS
Constructpercent	Percentage of employment in Construction Sector	1/1990-12/2018	calculated as $sector_i/totalemployment_i$
EDHLTH	Education and Healthcare employment	1/1990-12/2018	BLS
EDHLTHpercent	Percentage of employment in Education and Healthcare	1/1990-12/2018	calculated as $sector_i/totalemployment_i$
Government	Government employment	1/1990-12/2018	BLS
Govpercent	Percentage of employment in government	1/1990-12/2018	calculated as $sector_i/totalemployment_i$
Information	Information services employment	1/1990-12/2018	BLS
Infopercent	Percentage of employment in information sector	1/1990-12/2018	calculated as $sector_i/totalemployment_i$
Leihosp	Leisure and hospitality employment	1/1990-12/2018	BLS
Leihosppercent	Percentage of employment in Leihosp	1/1990-12/2018	calculated as $sector_i/totalemployment_i$
Manufacture	Manufacturing employment	1/1990-12/2018	BLS
Manufactpercent	Percentage of employment in manufacturing	1/1990-12/2018	calculated as $sector_i/totalemployment_i$
Other	Other services employment	1/1990-12/2018	BLS
Otherpercent	Percentage of employment in other services	1/1990-12/2018	calculated as $sector_i/totalemployment_i$
Professional	Professional and business services employment	1/1990-12/2018	BLS
Profpercent	Percentage of employment in Professional services	1/1990-12/2018	calculated as $sector_i/totalemployment_i$
Retail	Retail employment	1/1990-12/2018	BLS
Retailpercent	Percentage of employment in retail	1/1990-12/2018	calculated as $sector_i/totalemployment_i$
Totalemployment	Total non-farm employment on Long Island	1/1990-12/2018	BLS
Wholesale	Wholesale employment	1/1990-12/2018	BLS
Wholesalepercent	Percentage of employment in wholesale trade	1/1990-12/2018	calculated as $sector_i/totalemployment_i$

The results of the analysis (Tables 4-7) are suggestive that hurricanes and tropical storms have had some impacts on employment, and that Superstorm Sandy did have some long-lasting effects on sectoral employment on Long Island. The magnitude and direction of those impacts are however sectoral specific.

As can be seen in Tables 4-7, Sandy appears to have countered the effects of negative existing trends in some sectors such as manufacturing, wholesale and retail trade, professional employment, and construction. In other sectors such as leisure and hospitality employment and information services it does appear to reinforce existing trends. The regression results are suggestive that Superstorm Sandy did have some long-term impacts on the structure of Long Island employment.

Table 4: Total and Sectoral Impacts regressions

Total Employment

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(USRGDP)	0.699709	0.042027	16.64924	0
C	0.593781	0.391244	1.517674	0.1301
@TREND	-0.0007	0.000122	-5.716646	0
SANDY	0.03321	0.0056	5.930923	0
LOG(NYMINWAGE)	-0.028346	0.018753	-1.511551	0.1317
LITROPHUR	0.003242	0.004687	0.691727	0.4896

Construction

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(USRGDP)	2.847784	0.184578	15.42862	0
C	-22.51506	1.718321	-13.10294	0
@TREND	-0.00432	0.000537	-8.037232	0
SANDY	0.126205	0.024593	5.131771	0
LOG(NYMINWAGE)	0.107897	0.082361	1.310052	0.1911
LITROPHUR	0.03468	0.020585	1.684715	0.093

Education & Healthcare

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(USRGDP)	0.077608	0.033216	2.336479	0.0201
C	4.362525	0.309219	14.1082	0
@TREND	0.002027	9.67E-05	20.96089	0
SANDY	-0.011535	0.004426	-2.606497	0.0096
LOG(NYMINWAGE)	-0.091501	0.014821	-6.173608	0
LITROPHUR	-0.006689	0.003704	-1.805584	0.0719

Government

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(USRGDP)	0.206016	0.081978	2.513061	0.0125
C	3.281151	0.763173	4.299352	0
@TREND	0.000358	0.000239	1.499874	0.1346
SANDY	-0.086302	0.010923	-7.901212	0
LOG(NYMINWAGE)	-0.019017	0.03658	-0.519878	0.6035
LITROPHUR	-0.026113	0.009143	-2.85615	0.0046

Information

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(USRGDP)	1.585589	0.21481	7.381353	0
C	-10.73013	1.999765	-5.365697	0
@TREND	-0.003268	0.000626	-5.22374	0
SANDY	-0.110724	0.028621	-3.868652	0.0001
LOG(NYMINWAGE)	-0.283487	0.095851	-2.95758	0.0033
LITROPHUR	-0.010034	0.023957	-0.418843	0.6756

Leisure & Hospitality

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(USRGDP)	0.254987	0.218501	1.166984	0.2441
C	1.83631	2.034123	0.902753	0.3673
@TREND	0.000915	0.000636	1.437872	0.1515
SANDY	0.066361	0.029113	2.279449	0.0233
LOG(NYMINWAGE)	0.052409	0.097498	0.537544	0.5913
LITROPHUR	0.06957	0.024369	2.854884	0.0046

Table 5: Total and Sectoral Impacts regressions

		Manufacturing		
Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(USRGDP)	0.81328	0.125427	6.484086	0
C	-2.692104	1.167659	-2.305557	0.0218
@TREND	-0.003724	0.000365	-10.19707	0
SANDY	0.104602	0.016712	6.259202	0
LOG(NYMINWAGE)	0.033212	0.055967	0.593419	0.5533
LITROPHUR	0.000565	0.013988	0.040406	0.9678
		Other		
Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(USRGDP)	0.820314	0.061555	13.32654	0
C	-3.705772	0.573043	-6.466827	0
@TREND	9.92E-05	0.000179	0.553412	0.5804
SANDY	0.02156	0.008201	2.628818	0.009
LOG(NYMINWAGE)	-0.132076	0.027467	-4.808592	0
LITROPHUR	0.00495	0.006865	0.720984	0.4715
		Professional		
Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(USRGDP)	1.326674	0.089865	14.76303	0
C	-7.130019	0.836591	-8.522708	0
@TREND	-0.001165	0.000262	-4.453328	0
SANDY	0.06309	0.011973	5.269175	0
LOG(NYMINWAGE)	-0.173796	0.040099	-4.334187	0
LITROPHUR	0.012976	0.010022	1.294762	0.1964
		Retail		
Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(USRGDP)	0.895786	0.064191	13.95507	0
C	-3.145873	0.597581	-5.264349	0
@TREND	-0.001393	0.000187	-7.453524	0
SANDY	0.045962	0.008553	5.373965	0
LOG(NYMINWAGE)	-0.059464	0.028643	-2.07607	0.0387
LITROPHUR	-0.004583	0.007159	-0.640135	0.5226
		Wholesale		
Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(USRGDP)	1.020332	0.047195	21.61964	0
C	-5.151496	0.439357	-11.72509	0
@TREND	-0.002028	0.000137	-14.75342	0
SANDY	0.042909	0.006288	6.823723	0
LOG(NYMINWAGE)	0.009497	0.021059	0.450988	0.6523
LITROPHUR	0.000571	0.005263	0.108498	0.9137

Table 6: Percent of Total Employment regressions

Percent of Total Employment: Construction			
Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(USRGDP)	2.148075	0.153862	13.96108	0
C	-23.10884	1.432368	-16.13331	0
@TREND	-0.00362	0.000448	-8.080286	0
SANDY	0.092995	0.0205	4.536259	0
LOG(NYMINWAGE)	0.136243	0.068655	1.984458	0.0481
LITROPHUR	0.031438	0.01716	1.832105	0.0679

Percent of total employment: Education & Healthcare			
Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(USRGDP)	-0.622102	0.052069	-11.94754	0
C	3.768744	0.484738	7.7748	0
@TREND	0.002727	0.000152	17.98519	0
SANDY	-0.044746	0.006938	-6.449699	0
LOG(NYMINWAGE)	-0.063155	0.023234	-2.718192	0.0069
LITROPHUR	-0.009931	0.005807	-1.710108	0.0882

Percent of total employment: Government			
Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(USRGDP)	-0.493693	0.102857	-4.799784	0
C	2.68737	0.957546	2.806518	0.0053
@TREND	0.001058	0.0003	3.53118	0.0005
SANDY	-0.119513	0.013705	-8.72066	0
LOG(NYMINWAGE)	0.009329	0.045896	0.203258	0.8391
LITROPHUR	-0.029355	0.011471	-2.559012	0.011

Percent of total employment: Information			
Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(USRGDP)	0.88588	0.221279	4.003454	0.0001
C	-11.32391	2.059986	-5.497084	0
@TREND	-0.002568	0.000644	-3.985294	0.0001
SANDY	-0.143935	0.029483	-4.881992	0
LOG(NYMINWAGE)	-0.255142	0.098738	-2.584037	0.0102
LITROPHUR	-0.013276	0.024678	-0.537975	0.591

Percent of total Employment: Leisure & Hospitality			
Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(USRGDP)	-0.444723	0.199196	-2.232587	0.0263
C	1.242529	1.854407	0.670041	0.5033
@TREND	0.001615	0.00058	2.783321	0.0057
SANDY	0.03315	0.026541	1.249047	0.2126
LOG(NYMINWAGE)	0.080755	0.088884	0.908547	0.3643
LITROPHUR	0.066327	0.022216	2.985617	0.0031

Percent of total employment: Manufacturing			
Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(USRGDP)	0.113571	0.1073	1.058449	0.2907
C	-3.285885	0.998901	-3.289499	0.0011
@TREND	-0.003025	0.000312	-9.680728	0
SANDY	0.071392	0.014296	4.993662	0
LOG(NYMINWAGE)	0.061558	0.047879	1.285708	0.1995
LITROPHUR	-0.002677	0.011967	-0.223699	0.8231

Table 7: Percent of Total Employment regressions

Percent of total employment: Other				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(USRGDP)	0.120605	0.036084	3.342303	0.0009
C	-4.299553	0.335927	-12.79908	0
@TREND	0.000799	0.000105	7.602051	0
SANDY	-0.01165	0.004808	-2.423182	0.0159
LOG(NYMINWAGE)	-0.10373	0.016101	-6.442316	0
LITROPHUR	0.001707	0.004024	0.424263	0.6717
Percent of total employment: Professional				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(USRGDP)	0.626964	0.064721	9.6872	0
C	-7.7238	0.602517	-12.81923	0
@TREND	-0.000466	0.000188	-2.471317	0.014
SANDY	0.02988	0.008623	3.464975	0.0006
LOG(NYMINWAGE)	-0.14545	0.028879	-5.036467	0
LITROPHUR	0.009734	0.007218	1.348597	0.1784
Percent of total employment: Retail				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(USRGDP)	0.196076	0.050752	3.863395	0.0001
C	-3.739654	0.472477	-7.914998	0
@TREND	-0.000694	0.000148	-4.693307	0
SANDY	0.012751	0.006762	1.885679	0.0603
LOG(NYMINWAGE)	-0.031119	0.022646	-1.374107	0.1704
LITROPHUR	-0.007825	0.00566	-1.38243	0.1678
Percent of total employment: Wholesale				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(USRGDP)	0.320622	0.040242	7.967285	0
C	-5.745276	0.374634	-15.33569	0
@TREND	-0.001328	0.000117	-11.33214	0
SANDY	0.009698	0.005362	1.808726	0.0714
LOG(NYMINWAGE)	0.037843	0.017957	2.107467	0.0359
LITROPHUR	-0.002671	0.004488	-0.595152	0.5522

Conclusions

The analysis presented above is still preliminary and needs additional work. As was pointed out earlier, the analysis is indicative that Superstorm Sandy did have some long-term impacts that are manifesting themselves in various ways in the structure of the LI's economy. It is though, an open question as to how much Sandy has altered the preexisting trends already taking place on the region's economy. The increasing popularity of online shopping is dramatically altering the region's retail landscape, and this coincides with the period in which Sandy struck the island, and thus may also need to be instrumented for in the analysis. In summary, there is still some significant work to be completed on the analysis and additional evaluation of the data, before greater conclusions can be drawn.

References

Aerts, J.C.J.H. , W.J.W. Botzen, H. de Moel, and M. Bowman. 2013. "Cost estimates for flood resilience and protection strategies in New York City." *Annals of the New York Academy of Sciences* 1294: 1-104.

Blake, E.S., T.B. Kimberlain, R.J. Berg, J.P. Cangialosi and J.L. Beven II. 2013. "Tropical Cyclone Report Hurricane Sandy (AL182012) 22 – 29 October 2012". National Hurricane Center.

Faber, J.W. 2015. "Superstorm Sandy and the Demographics of Flood Risk in New York City." *Human Ecology* 43: 363–378.

Federal Reserve Bank of New York. 2018. "Long Island." https://www.newyorkfed.org/data-and-statistics/regional-data-center/profiles/longisland

FEMA. 2017. "Sandy Five Years Later." https://www.fema.gov/sandy-5-year.

Greenburg, M. 2014. "The Disaster inside the Disaster: Hurricane Sandy and Post-crisis Redevelopment." *New Labor Forum* 23(1): 44 –52.

Grube, L. E., R. Fike and V.H. Storr. 2017. "Navigating Disaster: An Empirical Study of Federal Assistance Following Hurricane Sandy." *Eastern Economic Jou*rnal 44: 576-593.

Luckman, M.H., D. Strafer, and C. Lipski. 2015. "Three Years Later, Sandy Survivors Remain Homeless." Touro Law Review 32(2): 313-337.

Neumann, J.E., K. Emanuel, S. Ravela, L. Ludwig, P. Kirshen, K. Bosma and J. Martinich. 2015. "Joint effects of storm surge and sea-level rise on US Coasts: new economic estimates of impacts, adaptation, and benefits of mitigation policy." *Climatic Change* 129: 337-349.

New York Rising. 2014. Community Reconstruction Committee Plans. https://stormrecovery.ny.gov/nyrcr/final-plans.

Ortega, F. and S. Taspinar. 2018. "Rising sea levels and sinking property values: Hurricane Sandy and New York's housing market." *Journal of Urban Economics* 106: 81–100.

Park, J., M. Son and C. Park. "Natural disasters and deterrence of economic innovation: a case of temporary job losses by Hurricane Sandy." *Journal of Open Innovation: Technology, Market, and Complexity* (2017) 3:5

Roy, Y. 2017. "Five years later, superstorm Sandy continues to change Long Islanders' lives." Newsday (Oct. 15). https://www.newsday.com/long-island/superstorm-sandy-five-years-later-1.14462179

Wanik, D.W., E. N. Anagnostou, M. Astitha, B. M. Hartman, G. M. Lackmann, J. Yang, D. Cerrai, J. He, and M. E. B. Frediani. 2018. "A Case Study on Power Outage Impacts from Future Hurricane Sandy Scenarios." *Journal of Applied Meteorology and Climatology* 57: 51-79.

Yang, B. and I. Jahan. 2018. "Comprehensive Assessment for Post-Disaster Recovery Process in a Tourist Town." *Sustainability* 10(6), 1842.

Revisit Lease Versus Purchase: Impact on Government Tax Revenue

Guan Jun Wang, Savannah State University

Abstract

Motivation for leasing is often believed to be the tax rate difference between the lessee and the lessor, allowing both save on taxes at the government's expense. This short note challenges this conventional wisdom and demonstrates the leasing impact on government tax revenue analytically and shows government is not always a loser.

Introduction

A lease is a contractual agreement under which the owner of an asset (the lessor) temporarily transfers the right to use of an asset to another party (the lessee). The lessor typically makes the lease for a specified time in return for periodic leasing payments from the lessee.

The advantages to lessees mentioned in leading managerial finance textbooks include higher operating flexibility, risk reduction and tax advantage: A firm that leases equipment or real estate, for example, will be able to deduct its lease payments from its taxable income immediately rather than deducting the cost of purchasing equipment as depreciation over time.

Lessors also benefit in several ways. Compared to lessees, they are often able to acquire equipment at a low cost, and obtain acceptable financing terms. These advantages result from various economies of scale, such as increased buying power with sellers of equipment. Lessors are also better positioned to take advantage of certain tax laws, such as depreciation allowances and investment tax credit.

Among these benefits, taxes are often believed to be the main reason that business choose to lease rather than purchase. A business purchasing an asset gains the tax benefit of depreciating the asset's cost. However, the depreciation is more valuable to firms in a higher tax bracket. Thus the conventional wisdom is that the business needing to use the asset lets a firm with a higher effective tax rate buy the asset and take the depreciation tax benefit. The asset's ultimate user then leases it from the purchaser under a contract that allows the gain to be shared between two businesses.

Many textbooks assert that because both the lessor and lessee gain from the transfer of depreciation from an asset's user to lessor who buys it, the government is a net loser. For example, Brealey, Myers, and Allen (2017) contend that "the mutual gain is at the expense of government". Ross, Westerfield, and Jordan (2019) state that "The loser will be the IRS" (P. 905). This conventional wisdom is also found in Brigham and daves (2019), Brigham and Ehrhardt (2017), and Emery, Finnerty, and Stowe (2018). The characterization of the lessee and lessor as paying lower aggregator taxes (which is equivalent to the government's collecting less) is not confined to textbooks. Myers, Dill, and Bautista (1976) observe that "saving taxes seems to be the only motive that is obvious and substantial, and Smith and Wakeman (1985) state "leasing can reduce total tax bill".

Review of lease versus purchase analysis

The modern standard conceptual framework of lease vs. purchase analysis is based on the notion that a lease contract is equivalent to debt. Since the asset user's operating cash flows are not affected by lease or purchase decision therefore may be ignored in a lease/purchase analysis. The net advantage to lease (NAL) is the difference between the investment outlay to buy the asset and the present value of the lease's net cash outflows (LNCOs), where each period's LNCO is the sum of the following three components: (1) the after-tax lease payment; (2) the foregone depreciation tax credit; and (3) the foregone interest tax credit on the debt that the lease is presumed to displace. If NAL>0, the asset user would prefer leasing to borrowing-to-buy.

A lessor's analysis of a proposed lease is a capital budgeting problem. The net present value (NPV) of the project is the difference between the present value of the lessor's net cash inflows (LNCI) and initial investment outlay. The lessor's net cash inflows in each period is the following three components: (1) the after-tax lease revenue; (2) the depreciation tax credit; (3) the interest tax credit on the debt. If NPV>0, lessor benefits from the leasing. The standard textbook approaches use after tax cost of debt as the discount rate. So both the net advantage to leasing (NAL) and the net present value the lessor obtains depend on lessee and lessor's tax rates and borrowing rates and the depreciation method, and of course lease contract stating the lease payment schedule.

Leasing impact on government tax revenue

The standard textbook analysis of a user's lease vs purchase decision emphasizes the tax-deductible amount: lease payment to the lessee and depreciation to the lessor (purchaser), and overlook the high cost, the lease payment, to the lessee and the revenue generated to the lessor from leasing. If the annual lease payment is higher than the annual depreciation which is usually a case, lessor's taxable income increases, so does the tax liability, thus government's tax revenue.

Let D_n denote depreciation in year n, L_n lease payment in year n, and I_{pn} buyer's interest payment in year n (if buyer uses 100% equity financing for the purchase, $I_{pn} = 0$), $I_{\ln}$ interest payment in year n incurred to the lessee if lessee had chosen to borrow and buy, Tb buyer (lessor)'s tax rate, T_l lessee's tax rate. Then the tax difference for user in year n between leasing and purchasing is

$$-(L_n - D_n - I_{\ln})T_l \tag{1}$$

Additional tax the lessor pay in year n (to the lessor leasing payment is the revenue) as a result of purchasing first then leasing out is

$$(L_n - D_n - I_{pn})T_p \tag{2}$$

The total effect on government's net tax revenue in year n (NTR_n) as a result of leasing is

$$NTR_n = (L_n - D_n - I_n)T_p - (L_n - D_n - I_{\ln})T_l \tag{3}$$

Whether government is a loser or gainer depends on both parties' tax rates and interest rates, depreciation method and leasing contract stating the payment schedule (amount and timing). If both parties have the same tax rate and borrowing rate and use debt financing the purchase, the government's net tax revenue in year n (NTR_n) is obviously zero, that is, leasing has no effect on government tax revenue. In many instances where the lessor is a large finance company while the lessee is a small industrial company, in such case, $T_p > T_l$ often holds and the borrowing rate of the lessee (the small industrial company) is often higher than that of the lessor (the large finance company), thus the sign and value of NTR_n not only depends on tax rate difference but also the borrowing rate, furthermore the depreciation method, leasing contract, The overall effect should be measured by the present value of the NTR_n in each period and discount rate used should be the risk free rate.

Equal tax rates and interest rates

When both parties (Lessor and Lessee) have the same tax rate and borrowing rate and use debt financing the purchase, that is $I_{pn} = I_{\ln}$ and $T_p = T_l$, then

$$NTR_n = (L_n - D_n - I_{pn})T_p - (L_n - D_n - I_{\ln})T_l = 0 \tag{4}$$

the government's net tax revenue in year n (NTR_n) is zero, that is, leasing has no effect on government tax revenue.

Unequal tax rates, straight line depreciation

The lessor is typically a large finance company while the lessee is a small industrial company, in such instance, the following is usually true:

$T_p > T_l$

$I_{pn} < I_{ln}$ (larger companies usually pay lower borrowing rates)

$L_n > D_n + I_n$ (lease payments are usually accelerated relative to the asset's depreciation and interest schedule).

Thus $NTR_n > 0$, which means Government can also gain from the leasing arrangement which is contradict to the conventional wisdom that the government is a net loser.

Unequal tax rates, accelerated depreciation

Under accelerated depreciation assumption, depends on the lease payment contract, the early years' NTR can be negative, and the present value of all NTR_n can be either positive or negative depending the leasing payment schedule and the risk free rate (the discount rate), so government can be either a loser or gainer.

Numerical examples

The numerical example in Ross, Westerfield, Jaffe and Jordan (2019), where the user's tax rate is assumed to be 0% and the lessor's is 21%. The lessor and lessee are assumed to have the same borrowing rate, 6.329%. The asset's outlay is $10,000, which is depreciable by the straight –line method for 5 year. The lease contract stipulates five annual lease payments of $2,393, due at the end of each year. Since the user's tax rate is 0%, leasing or purchase doesn't affect user's tax position. And lessor's additional tax liability = $(L_n - D_n - I_{pn})T_p$ is negative under the above assumption, in this case, leasing reduces lessor's tax liability and government loose revenue. In this example, the low lease payment (Ln <Dn+ I pn) results in government loss of revenue. However, a lease payment typically includes the sum of the following three components: depreciation, interest charge and sale's tax, therefore annual lease payments of $2,393 while depreciation is $2000 and interest payment is over $600 is an unrealistic assumption.

Holding all else the same, if we change lease payment to higher amount at least covering depreciation and interest, the net government tax revenue will be positive. All else the same, government tax revenue would be lower with accelerated depreciation, and higher with accelerated lease payment.

Conclusion

This short notes discusses the leasing impact on government tax revenue both analytically and numerously and shows government is not always a loser, which contradict to a widely held view "Government is a net loser in leasing arrangement".

This short note provides insights that are not well known and should be useful to both instructors and practitioners.

References

Brealey, R., Myers, S. and Allen, F.,2017. *Principles of Corporate Finance*, 10th Edition, McGraw-Hill Irwin

Brighham, E., and Daves, P., 2019, *Intermediate Financial Management*, 13th Edition, Cengage Learning

Brigham, E. , and Ehrhardt, M., 2017, *Financial Management: Theory and Practice*, 15th Edition, Cengage Learning

Emery,D., Finnerty, J., and Stowe, J., 2018, *Corporate Financial Management*, 5th Edition, Wohl Publishing

Myers,S., Dill, D., and Bautista, A., 1976, Valuation of Financial Lease Contracts, *Journal of Finance*, 31 (3): 799-819

Ross, S., Westerfield,R., and Jordan, B., 2019*, Fundamentals of Corporate Finance*, 13th Edition, McGraw-Hill Irwin

Ross, S., Westerfield, R., and Jordan, B., 2019, *Corporate Finance*, 13th Edition, McGraw-Hill Irwin

Smith, C., and Wakeman, L., 1985. "Determinants of Corporate Leasing Policy." *Journal of Finance*, 40 (3): 895-908

Expanding the Scope of the Piotroski Model
Melissa K. Woodley, Creighton University
Steven T. Jones and James P. Reburn, Samford University

Abstract

The Piotroski (2000) model was initially designed as a means of differentiating future winners from future losers among value stocks. Piotroski confined the tests of his model to value stocks because of his belief that growth stocks were less likely than value stocks to prove amenable to fundamental analysis.

In this paper, we explore whether the effectiveness of the Piotroski accounting screens might in fact extend to growth stocks in addition to value stocks. We do this by applying Piotroski's tests to all five book-to-market quintiles, rather than restricting our sample to the value quintile. The relative responses of the various book-to-market quintiles to the Piotroski screens vary across time periods. However, in no case do we see evidence that the model's efficacy is strictly limited to the "value" quintile. Further, it is often the case that this is not even the quintile for which the model's performance is strongest.

Introduction, Literature Review, and Motivation

Value Stocks as Particular Targets for Fundamental Analysis

Piotroski (2000) points out that several papers prior to his own have looked for ways to identify stocks whose value the market has underestimated as a result of overly pessimistic expectations. In particular, Frankel and Lee (1998), Dechow and Sloan (1997), and LaPorta (1996) have explored this issue.

Piotroski notes two potential reasons for viewing fundamental analysis as particularly promising with regard to value stocks. One reason is that to the extent that fundamental analysis "works" in evaluating value stocks, it may enable the investor to realize especially high return levels. The reason for this is that while value stocks as a group earn above-average returns, the majority of individual value stocks actually produce below-average returns. This result, of course, can only occur if the average outperformance of those value stocks that do outperform the market outweighs the average underperformance of those value stocks that underperform the market; in other words, taken as a group those value stocks that outperform the market apparently do so by a fairly wide margin.

Thus, if one can find a way to isolate those individual value stocks that are likely to provide above-average returns, the overall return advantage that one could obtain by investing in value stocks can, at least in theory, be magnified considerably. Holthausen and Larcker (1992), Lev and Thiagarajan (1993), and Abarbanell and Bushee (1998) have all made strides in using fundamental analysis to estimate future market performance.

In addition, Piotroski argues that value stocks are more likely than growth stocks to lend themselves to fundamental analysis to begin with. He explains this as follows: "From a fundamental analysis perspective, value stocks are inherently more conducive to financial statement analysis than growth (i.e., glamour) stocks. Growth stock valuations are typically based on long-term forecasts of sales and the resultant cash flows, with most investors relying heavily on nonfinancial information. Moreover, most of the predictability in growth stock returns appears to be momentum driven (Asness [1997]). In contrast, the valuation of value stocks should focus on recent changes in firm fundamentals (e.g., financial leverage, liquidity, profitability, and cash flow adequacy) and an assessment of these characteristics is most readily accomplished through a careful study of historical financial statements. To the extent that investors can use financial statement analysis to identify strong value companies, a firm-specific, high-return investment strategy based on the BM effect can be created." (Piotroski 2000, p. 5.)

It is this contention – the notion that value stocks lend themselves more readily than do growth stocks to fundamental analysis of the type explored by Piotroski (2000) – that we explore in detail in this paper. Is Piotroski correct in assuming that one might as well not attempt to apply his methodology to the selection of growth stocks? Or, is it possible that he is being too modest in assuming that the usefulness of his stock-selection model is limited to value stocks?

The Piotroski Methodology

As described in our previous work on this topic (Woodley, Jones, and Reburn, 2011): "Since the present paper is intended primarily as an attempt to replicate Piotroski's (2000) results, we will describe his work in somewhat more detail than would ordinarily be included in a literature review. For each year from 1976 through 1996, Piotroski identifies those firms whose book-to-market ratios fall into the highest quintile. To expand on his basic results, he performs a separate division of firms into

terciles, based on market capitalization. His set of high book-to-market stocks is then subdivided based on whether these stocks fall into the high, medium, or low market capitalization tercile of the overall market.

"Each stock in the top book-to-market quintile is then evaluated on nine separate factors, which we itemize below, and receives a score of either 1 ("good") or 0 ("bad") on each of these factors. The firm's scores on these 9 factors are summed, resulting in an "F_Score" ranging from 0 to 9, inclusive. Firms that have higher F_Scores are hypothesized to be the most likely to produce positive market-adjusted returns over the ensuing year, and vice versa. Market-adjusted return realizations are evaluated separately for firms with each score from 0 through 9; in addition, results are evaluated for firms with scores of 0 and 1 combined ("Low Score") and firms with scores of 8 and 9 combined ("High Score").

"The nine factors that Piotroski (2000) considers can be divided into indicators of the following three general attributes: profitability; leverage, liquidity, and source of funds; and operating efficiency. In the area of profitability, four specific indicators are chosen. Scores of 1 are assigned for each of the following outcomes: ROA (net income before extraordinary items over beginning-of-year total assets) is positive; CFO (cash flow from operations over beginning-of-year total assets) is positive; ΔROA (current year's ROA minus prior year's ROA) is positive; and ACCRUAL (ROA minus CFO) is negative. Otherwise, scores of 0 are assigned for the respective factors.

"In the area of leverage, liquidity, and source of funds, three specific indicators are chosen. Scores of 1 are assigned for each of the following outcomes: ΔLEVER (the most recent year's ratio of long-term debt to average total assets, minus the corresponding ratio for the prior year) is negative; ΔLIQUID (the most recent year's ratio of current assets to current liabilities, minus the corresponding ratio for the prior year) is positive; and EQ_OFFER (an issuance of common equity within the past year) did not occur.

"In the area of operating efficiency, two specific indicators are chosen. Scores of 1 are assigned for the following: ΔMARGIN (current year's ratio of gross margin to total sales, minus the corresponding number for the prior year) is positive; and ΔTURN (current year's ratio of total sales to beginning-of-year total assets, minus the corresponding number for the prior year) is positive."

Piotroski's Findings, and Ours

As shown in his Table 3, Piotroski (2000) finds a fairly consistent pattern in which market-adjusted returns increase as the F_Score increases. Further, this superiority of returns for "High Score" firms is highly statistically significant across a number of comparisons. In his Table 4, Piotroski shows that the outperformance of "High Score" firms is extremely strong among small firms, is also quite impressive among mid-size firms, and is rather weak among large firms.

Our own results (Woodley, Jones, and Reburn, 2011) for the Piotroski (2000) test period of 1976-1996 are qualitatively similar to Piotroski's, confirming that "High Score" firms tend to produce above-average returns. However, during a subsequent test period from 1997-2008, our findings are exactly reversed from those during the Piotroski test period. On average, "High Score" firms underperform value stocks as a group by a wide margin, and underperform "Low Score" firms by an even wider margin. Further, we find that both our agreement with Piotroski's findings that "High Score" firms outperform over his test period, and our result that "High Score" firms underperform over the ensuing 12-year period, are largely invariant to firm size. While these results by no means prove that the Piotroski model has lost its usefulness, they would certainly be consistent with a scenario in which a good model has gone bad.

In subsequent work (Woodley, Jones, and Reburn, 2019 (forthcoming)), we find encouraging news for those who would like to continue using the Piotroski (2000) screens as a means of selecting those value stocks that are most likely to perform well. Our results here find that during a now longer overall test period, the Piotroski screens produce fairly good results. More detailed testing indicates that the ability of the Piotroski model to find value stocks that will outperform seems to vary based primarily on investor sentiment. Specifically, at those times when the volatility index ("VIX," also colloquially known as the "fear index") is relatively low, the model works very well. Results are noticeably weaker during times when the VIX is at more average levels, and actually works in reverse during time periods in which the VIX is high. Thus, the seeming weakening of the model that was identified in our prior work appears to have been related more to the relative levels of investor nervousness during the test periods in question, rather than a weakening of the model over time.

Motivation: Assessing the Assumption that the Piotroski Methodology's Effectiveness is Limited to Value Stocks

Piotroski's (2000) results regarding the usefulness of fundamental analysis in the selection of value stocks are, quite clearly, both interesting and (at least potentially) quite useful. Given this usefulness, it would seem to be worthwhile to examine whether the model can be applied to other stocks, rather than simply assuming that the benefits of the Piotroski methodology are limited to value stocks.

Data and Methodology

As described in our previous work on this topic (Woodley, Jones, and Reburn, 2011): "Using financial statement data from Compustat, and market returns and market capitalization data from CRSP, the following methodology is employed for each fiscal year in the sample period (1976-2008). For each fiscal year T, each firm's book-to-market ratio and total market value are calculated as of the fiscal year end date for fiscal year T-1. (See Piotroski 2000, p. 11, footnote 8.) Firms are sorted into quintiles based on their book-to-market ratios, and are separately sorted into terciles based on size. Each firm that falls within the top book-to-market quintile is considered part of the sample of value firms, subject to availability of all necessary financial data and market return data.

"For each such firm, each of the financial indicators described above is calculated for fiscal year T, and the firm's F_Score for fiscal year T is calculated based on these indicators. Raw returns and market-adjusted returns are then calculated for the one-year period beginning in the fifth month after the end of fiscal year T. For instance, for a firm using a calendar year, the 2008 fiscal year ends on December 31, 2008; the corresponding returns are calculated for the one-year period beginning May 1, 2009 and ending April 30, 2010. CRSP returns data are currently available only through 2010, thus necessitating the choice of 2008 as the last date for "year T." This is also why Piotroski's (2000) final "year T" is 1996.

"An observation is dropped from the sample if the firm's fiscal year end date for fiscal year T is not clear in Compustat, if the firm's fiscal year T lasts for a period other than 12 months (due to a change in fiscal year end date from one year to the next), or if there is not sufficient information to calculate all variables of interest, including those that involve changes from fiscal year T-1.

"This process is repeated for each year from 1976-2008. All observations with a given F_Score, regardless of the specific year within the sample period, are initially grouped together for purposes of determining the distribution of returns for that F_Score."

In this paper, we make three key adjustments to the methodology described above. First, as noted above, the purpose of this paper is to learn whether the efficacy of the Piotroski (2000) model is limited to "value stocks." So, while we continue to sort the firms into book-to-market quintiles, we do not limit our testing to those stocks that fall into the highest book-to-market quintile (the "value" quintile).

Second, due to the passage of time, we are able here to use a longer sample period than that described in Woodley, Jones, and Reburn (2011) or even in Woodley, Jones, and Reburn (2019, forthcoming). In those papers, the last years for "year t" were 2008 and 2013, respectively. Our final "year t" for purposes of this paper is 2015, because 2017 is the last year for which full-year returns data are available in the latest (2018) CRSP data release. Based on the timeline described above, for a firm whose reporting period ends on December 31, 2015, we calculate the ensuing one year's returns from May 1, 2016 through April 30, 2017. Thus, in this paper the sample includes fiscal years 1976-2015, inclusive.

Third, we add a new set of tests. In the Piotroski (2000) papers, and in each of our previous papers, for every sample or sub-sample that is tested there are four sets of tests. Both raw returns and market-adjusted returns are tested, and in each case the tests compare the returns of "High F_Score" firms to both the returns of the overall sample, and the returns of "Low F_Score" firms. In each of these four sets of tests, three specific comparisons are evaluated: mean return, median return, and percentage of positive returns. So, each sample or sub-sample is subjected to 12 specific tests.

We add two more sets of tests, in which returns are compared for High F_Score firms vs. all firms, and for High F_Score firms vs. Low F_Score firms, with the returns in question being adjusted based on Carhart's (1997) extension of the Fama-French three-factor model to include a fourth factor based on momentum. Here, too, each set of tests compares mean returns, median returns, and percentage of positive returns. Thus, for each sample or sub-sample that is tested, there are a total of 18 specific tests of comparative returns.

We evaluate the statistical significance of mean returns using a t-test, the statistical significance of median returns using a median two-sample test (with a two-sided test of the significance of the Z-statistic), and the percentage of positive returns using a binomial proportion difference with Wald confidence intervals (again with a two-sided test of the significance of the Z-statistic). Thus, for each sample or sub-sample that is tested, there are a total of 18 specific tests.

Results

For purposes of brevity, in this proceedings version of the paper the tables described in this section are omitted.

Results Across Entire Universe of Stocks

Table 1 shows our results for the entire sample, without regard to book-to-market quintile. Results are first determined over the full test period (fiscal years 1976-2015). In each of the 18 tests described above, use of the Piotroski (2000) screen produces

superior returns, and in each case the p-value is below 0.0001. During the Piotroski test period of 1976-1996, we again find that in all 18 tests, the model produces superior returns, with a p-value below 0.0001.

During the "post-Piotroski" test period, we again find that in each of the 18 statistical comparisons the results are in the expected direction. In the case of both the median return and the percentage of positive returns, all results are yet again significant with p-values of below 0.0001. However, only one of the six tests of the mean return produces results that are statistical significant at a level of 0.10 or better. This one result occurs when we compare the 4-factor model adjusted returns for High F_Score firms to those of the entire sample; here, the result is an impressive p-value of 0.0234.

So, the model performs extremely well in the overall test period and in both sub-periods, if the point of comparison is either median return or percentage of positive returns. If the point of comparison is mean return, the results are always in the predicted direction. These results are consistently, highly significant for both the overall test period and the first sub-period, but generally are not statistically significant during the second sub-period.

Results By Book-to-Market Quintile: Full Test Period

It would seem unlikely that results as strong as those described above would obtain for the entire universe of stocks, if the Piotroski (2000) screen were effective only for the highest book-to-market quintile. However, the observations described above do not definitively prove that the model works within any particular quintile, and do not even begin to address the relative levels of efficacy of the model across the various quintiles. So, we proceed to test the effectiveness of the model within each of the quintiles.

Table 2 displays the results of these tests when they are applied to our entire sample period of 1976-2015. These results can be summarized as follows.

First, the lowest book-to-market ("growth") quintile produces consistently strong results for the Piotroski (2000) screen. All 18 comparisons are in the expected direction, and all are significant at the 0.01 level. 15 of the 18 have p-values below 0.0001. The 3 exceptions are the three comparisons of mean results for High F_Score stocks vs. Low F_Score stocks," and in these three cases the "absolute" numbers are impressive, with the weakest p-value being 0.0017. Clearly, over the test period as a whole the model works exceedingly well for the growth quintile.

While the results are slightly weaker for the second quintile, they remain impressive. All comparisons are in the expected direction, and the weakest p-value is 0.0158. All other results are easily significant at the 0.01 level. There is a bit of further weakening when we move to the third quintile. In all 12 of the comparisons that were run by Piotroski (2000) and in our own prior work, results are still quite good, with the weakest p-value being 0.0135. With the tests of the 4-factor model, 5 of the 6 tests are highly significant. However, the difference in mean returns for the High F_Score group and the Low F_Score group, while still in the expected direction, is no longer significant, with a p-value of 0.2374. Regarding the fourth quintile, in the 12 "traditional" tests there is slight further weakening relative to the third quintile. In tests of the 4-factor model, we consistently find results that are significant at the 0.10 level, with the highest p-value being 0.0570. Thus, the general trend within the middle three quintiles is that the model remains very strong, but tends to weaken slightly as we move further away from the growth quintile.

Finally, we test the highest book-market ("value") quintile. For the 12 "traditional" tests, all results are still in the expected direction, and all results are significant at the 0.10 level. 11 of the 12 test results are significant at the 0.05 level, the sole exception being the comparison of mean market-adjusted returns for the High F_Score group vs. the entire sample. With the 6 tests of the 4-factor adjusted returns, all results are in the expected direction. The difference in the mean returns for the High F_Score group and the overall sample is not significant (p-value = 0.2288), whereas the difference in the mean returns for the High F_Score group and the Low F_Score group is highly significant (p-value = 0.0009). All other p-values are below 0.0001.

On balance, the screen performs best in the first BM quintile (lowest BM, i.e., "growth"), and it performs nearly as well in the second BM quintile. Model performance is somewhat weaker in quintiles 3, 4, and 5 than it is in quintiles 1 and 2. Thus, while the Piotroski (2000) model was originally created with value stocks in mind, for our overall test period of 1976-2015 it performs at least as well for growth stocks as for value stocks. Indeed, to the extent that there is a difference in the performance of the model, it actually performs better for growth stocks than for value stocks.

Results By Book-to-Market Quintile: Piotroski Test Period (1976-1996)

As noted above in our discussion of Table 1, the results for the overall sample were noticeably stronger during the 1976-1996 test period corresponding to the test period examined in Piotroski (2000) than in the subsequent test period of 1997-2015. Table 3 performs the same comparisons across quintiles as those shown in Table 2, but does so specifically during the "Piotroski window."

We again begin by performing tests of the model's effectiveness when applied within the lowest book-to-market ("growth") quintile. All 18 comparisons produce results in the expected direction, and in every case the results are highly statistically significant, with each p-value below 0.0001.

In tests of the second quintile, the only result that is not significant at the 0.05 level is the comparison of median returns, when returns are adjusted for the 4-factor model, and the comparison is between the returns of the High F_Score group and the overall sample. Even in this one instance, the p-value is an impressive 0.0523. Overall test results for the second quintile are very strong, although not as strong as for the first quintile.

The results for the third quintile are in some instances a bit stronger, and in others a bit weaker, than for the second quintile. We do find one result that is not statistically significant: specifically, in comparing the 4-factor-adjusted returns for High F_Score vs. Low F_Score firms, the p-value for the difference in means is 0.1356. Likewise, for the fourth quintile, when we compare the 4-factor-adjusted returns for High F_Score vs. Low F_Score firms, the difference in means is not statistically significant, with a p-value of 0.1321. All other results are significant at the 0.05 level or better.

When we turn to the highest book-market ("value") quintile, the results strengthen relative to the three middle quintiles. The weakest p-value among the 18 tests is 0.0122. This result is for the comparison of 4-factor adjusted returns, when the point of comparison is the mean return for the High F_Score group vs. that of the entire sample. In the 12 "traditional" comparisons, all p-values are below 0.0001.

So, for the 1976-1996 time period tested by Piotroski (2000), the results are very strong across all five quintiles, but somewhat stronger at either extreme than in the middle. Forced to choose, we would probably say that the growth quintile's results are slightly stronger than those of the value quintile; but, we would readily admit that the difference here is quite small. The key point, relative to the issue that our research is designed to address, is that the growth quintile's results are, at minimum, equally as strong as those of the value quintile.

Results By Book-to-Market Quintile: Post-Piotroski Test Period (1997-2015)

Table 4 displays the results of the same quintile-by-quintile tests as those displayed in Tables 2 and 3, but for our "post-Piotroski" test period of 1997-2015. Beginning as always with the lowest book-market ("growth") quintile, we find – broadly consistent with our Table 1 findings regarding the entire sample – that results are much weaker during this period than they are for Piotroski's (2000) test period of 1976-1996, and also much weaker during this period than they are for our overall test period of 1976-2015.

This is not to say that the model falls apart completely. In comparing the returns for the High F_Score firms to those for the sample as a whole, all results are in the expected direction, and all nine statistical tests are significant at the 0.05 level, with seven results significant at the 0.01 level and with five of those having p-values below 0.0001. However, when the returns of High F_Score firms are compared to those of Low F_Score firms, none of the results are remotely significant, and in fact only two of the nine have the expected sign.

Results for the second quintile are arguably a bit stronger, on the whole, than those for the first quintile. For the third quintile, results when comparing High F_Score firms to those of the entire sample are a bit weaker than they were for the two more "growth" oriented quintiles; but, results when comparing High F_Score firms to those of Low F_Score firms are much better than the results for the two more "growth" oriented quintiles. For the fourth quintile, none of the 6 tests of mean returns produce a statistically significant result, although 5 of the 6 do have the expected sign. However, all tests for median returns, and all tests for the percentage of positive returns, have the expected sign; further, all are significant at the 0.05 level, and 10 of the 12 are significant at the 0.01 level.

Finally, we turn to the highest book to market ("value") quintile. As with the fourth quintile, none of the tests of mean returns produce a statistically significant result; unlike the fourth quintile, only 3 of these 6 tests produce the expected sign. Like the fourth quintile, all tests of median returns and market-adjusted returns produce the expected sign. 10 of these 12 results are significant at the 0.10 level, 9 are significant at the 0.05 level, and 6 are significant at the 0.01 level.

Thus, results for the value quintile are clearly stronger than those for the first two quintiles, but are somewhat weaker than those for the fourth quintile. How the results for the value quintile compare to those for the third quintile is arguable. Clearly, the overall trend during this period is for the model to perform better as we move from the "growth" end of the spectrum toward the "value" end of the spectrum. But, even here the model's usefulness is not confined to the highest book-to-market quintile, and indeed this is not even the quintile in which the model's overall performance is strongest.

Conclusions

Piotroski's (2000) model was initially developed specifically for distinguishing future winners from future losers among value stocks, and Piotroski offers what we would consider a highly plausible argument as to why value stocks might be

particularly well-suited for the type of fundamental analysis that the model entails. However, it is surprising that, at least so far as we are aware, the assumption that the model would be suitable only for this subset of stocks has never been tested directly.

We find that during an expanded test period of 1976-2015, the model is actually at its strongest in the lowest book-to-market quintile; i.e., the "growth stock" quintile. Further, the model's second-strongest results during this period are in the second lowest book-to-market quintile.

When we test the model during Piotroski's (2000) original test period of 1976-1996, we do find that the model works better for the highest book-to-market ("value") quintile than it does for any of the three middle quintiles. However, during this period the model works at least as well for growth stocks as it does for value stocks, and if anything may actually work a bit better for growth stocks.

During our subsequent period of 1997-2015 – the period in which the model's overall performance is the weakest – it is true that the model's performance tends to improve as we move from the "growth" end of the spectrum to the "value" end of the spectrum. But, even here, the model's effectiveness is not confined to the highest book-to-market quintile, and in fact this is not even the single quintile within which the model's performance is strongest – that distinction belongs to the book-to-market quintile immediately below the value quintile.

So, in no case do we see evidence that the model's usefulness is confined to stocks at the value end of the growth-value spectrum. This would seem to imply that Piotroski (2000) may indeed have been overly modest regarding the potential applicability of his model.

References

Abarbanell, Jeffery S., and Brian J. Bushee. Spring 1997. "Fundamental Analysis, Future Earnings, and Stock Prices." *Journal of Accounting Research* 35(1): 1-24.

Asness, Clifford S. March/April 1997. "The Interaction of Value and Momentum Strategies." *Financial Analysts Journal* 53(2): 29-36.

Carhart, Mark M. March 1997. "On Persistence in Mutual Fund Performance." *Journal of Finance* 52 (1): 57-82.

Dechow, Patricia M., and Richard G. Sloan. January 1997. "Returns to Contrarian Investment Strategies: Tests of Naïve Expectations Hypotheses." *Journal of Financial Economics* 43(1): 3-27.

Holthausen, Robert W., and David F. Larcker. June-September 1992. "The Prediction of Stock Returns Using Financial Statement Information." *Journal of Accounting and Economics* 15(2-3): 373-411.

LaPorta, Rafael. December 1996. "Expectations and the Cross-Section of Stock Returns." *Journal of Finance* 51(5): 1715-42.

Lev, Baruch, and S. Ramu Thiagarajan. Autumn 1993. "Fundamental Information Analysis." *Journal of Accounting Research* 31(2): 190-215.

Piotroski, Joseph D. 2000. "Value Investing: The Use of Historical Financial Statement Information to Separate Winners from Losers." *Journal of Accounting Research* 38(Supplement): 1-41.

Woodley, Melissa K., Steven T. Jones, and James P. Reburn. December 2011. "Value Stocks and Accounting Screens: Has a Good Rule Gone Bad?" *Journal of Accounting and Finance* 11(4): 87-104.

Woodley, Melissa K., Steven T. Jones, and James P. Reburn. 2019 (forthcoming). "The 'Fear Index' and the Effectiveness of the Piotroski Model." *Quarterly Journal of Finance and Accounting.*