FINANCIAL SERVICES REVIEW Volume 28 Number 1 (2020)

Volume 28, No. 1
2020

ISSN 1057-0810

28 (1) 1–65
(2020)

FINANCIAL SERVICES REVIEW

The Journal of
Individual Financial Management

CONTENTS

AFS
Education • Research • Practice
Academy of Financial Services

FPA
FINANCIAL PLANNING ASSOCIATION

Official Journal of the Academy of Financial Services
published with the Financial Planning Association

ACADEMY OF FINANCIAL SERVICES

published in collaboration with the
Financial Planning Association

Financial Services Review is the journal of the Academy of Financial Services, published in collaboration with the Financial Planning Association. Membership dues of $125 to the Academy include a one-year subscription to the journal. Financial Planning Association members receive digital access to the current volume/issue of the journal.

How to Submit: Membership in AFS ($125) is required to submit an article to Financial Services Review. Join AFS at academyfinancial.org. A submission fee of $100 per article should be paid at: *https://academyoffinancialservices.wildapricot.org/Submit-an-Article*. Submit your article electronically as an email attachment in WORD format only (no PDFs please) to the Editor Stuart Michelson at *smichels@stetson.edu*. Should a manuscript revision be invited, no additional fees will be required. Style information for the manuscripts can be found on the inside back cover of this journal.

Financial Services Review

The Journal of Individual Financial Management

Vol. 28, No. 1, 2020

Co-Published by the Academy of Financial Services and the Financial Planning Association

The editor of *Financial Services Review* wishes to thank the Stetson University, School of Business, for its continuing financial and intellectual support of the journal.

Aims and Scope: *Financial Services Review* is the official publication of the Academy of Financial Services. The purpose of this refereed academic journal is to encourage rigorous empirical research that examines **individual behavior** in terms of financial planning and services. In contrast to the many corporate or institutional journals that are available in finance, the focus of this journal is on individual financial issues. The Journal provides a forum for those who are interested in the individual perspective on issues in the areas of Financial Services, Employee Benefits, Estate and Tax Planning, Financial Counseling, Financial Planning, Insurance, Investments, Mutual Funds, Pension and Retirement Planning, and Real Estate.

Publication information. *Financial Services Review* is co-published quarterly by the Academy of Financial Services, and the Financial Planning Association. Institutional subscription price is $100. Personal subscription price is $125 and is available by joining the Academy of Financial Services. Further information on this journal and the Academy of Financial Services is available from the website, http://www.academyfinancial.org. Postmaster and subscribers should send change of address notices to Stuart Michelson, Academy of Financial Services, Stetson University, School of Business, 421 N. Woodland Blvd., Unit 8398, DeLand, FL 32723.

Editorial Office: Stuart Michelson, School of Business, Stetson University, 421 N. Woodland Blvd., Unit 8398, DeLand, FL 32723. Phone: (386) 822-7376. Email address: *smichels@stetson.edu*. Web address: www.academy financial.org.

Advertising information. Those interested in advertising in the journal should contact Stuart Michelson, School of Business, Stetson University, 421 N. Woodland Blvd., Unit 8398, DeLand, FL 32723. Email address: smichels@stetson.edu, (386) 822-7376.

Printed in the USA

FINANCIAL SERVICES REVIEW

Financial Services Review 28 (2020) v–vii

From the Editor

This issue contains **Volume 28 - Issue 1** of *Financial Services Review* (*FSR*). I would like to thank the board and members of the Academy of Financial Services for their continued support. I continue to work in broadening the scope of articles, while still focusing on individual financial management and personal financial planning. I encourage authors to reach out when discussing implications of their findings in a more comprehensive way. As such, all articles in the Journal more appropriately relate to financial planning issues.

The lead article "The decrease in life insurance ownership: Implications for financial planning" is coauthored by Kyoung Tae Kim at University of Alabama, Travis P. Mountain at Virginia Tech University, Sherman D. Hanna at Ohio State University, and Namhoon Kim at Korea Rural Economic Institute. Using the Survey of Consumer Finances dataset the authors find the proportion of households owning a life insurance policy decreased from 72% in 1992 to 60% in 2016. They estimate logistic regressions on the likelihood of ownership of term and cash value life insurance. They find that changes in household characteristics accounted for the decrease in term life insurance ownership, but not for the decreases in cash value life insurance ownership. They also find a positive association between use of a financial planner and life insurance ownership.

The second article "Are 'Fun' Sources of Windfalls Destined to be Spent Hedonistically?" is coauthored by Eugene Bland at Texas A&M University – Corpus Christi and Valrie Chambers at Stetson University. The authors show that fun sources of income are more likely to be spent on a fun expenditures. Money won on a game show would be spent more on 'fun' than money received from a tax rebate. They find support for rejecting the hypothesis that there is no difference in allocations for regular expenses, credit card payments, durable assets or investing in stocks, bonds and savings account ("adult" uses of funds) by source of windfall. They found significant evidence that there is a difference in investing based on the source of the windfall. People apparently spend significantly more on fun when a fun windfall is received, but that spending on fun is not limitless. Additionally they find that there may be such a thing as "enough spending on fun.

The third article, "A Portfolio of Leveraged Exchange Traded Funds" is coauthored by William J. Trainor Jr., Indudeep Chhachhi, and Christopher L. Brown, all at Western Kentucky University. In this study, the authors demonstrate how a portfolio of leveraged exchange traded funds (LETFs) outperforms a portfolio using traditional ETFs while simultaneously reducing downside risk. Their results are primarily a function of LETFs borrowing short while the investor lends the additional wealth generated from this leverage in 1 to

7- year Treasury bonds or similar type of assets. They also present that for every 1% earned above the implied borrowing rate, a portfolio of 2x and 3x LETFs outperforms a traditional portfolio by 0.41% and 0.63% respectively, They show that more than 90% of LETFs outperformance is explained by the borrowing lending differential.

The final article, "Are Multiple Share Class Funds Poorly Governed?" is coauthored by Jonathan Handy at Furman University and Thomas Smythe at Florida Gulf Coast University. Utilizing independent Morningstar Stewardship Grades, the authors find that multiple share class mutual funds (MS funds) have lower quality governance. Using ordered probit regressions they find that MS funds are more likely to have lower board quality ratings and managerial incentive ratings. Their results show that less sophisticated investors seeking financial advice (those typically utilizing MS funds) may potentially be directed to funds that underperform and have higher costs.

Thanks to those who make the journal possible, especially the referees and contributing authors. Over the past year, the following reviewers provided excellent reviews of the articles you enjoyed within the pages of Financial Services Review. I would like to send a special thank you to the many reviewers that have significantly contributed to the quality of our journal by providing timely and thorough reviews of the submissions to our journal.

John E. Grable	Tim Kaiser
Sina Ehsani	Travis L. Jones
Michael Highfield	Swarn Chatterjee
William Trainor	Sarah Reiter
John Clinebell	Shaun Pfeiffer
Sophie Shive	Yuliya Plyakha
Xu Sun	Gene Stout
Timothy Krause	John Clinebell
Laszlo Sandor	Anders Carlander
Lewis W. Coopersmith	Kenneth Ryack
Jerry Stevens	Sonya Britt
Robert Ottr	Gary Porter
Kent Baker	David Yeske
Jean Lown	Scott Moore
Stuart Heckman	Jonathan Guyton
Haiwei Chen	Shibashish Chakraborty
Wade D. Pfau	Jiri Sindelar
Timothy Lu	Tianyang Wang
Kaustav Misra	Daniel Huerta-Sanchez
Sandeep Singh	Sara Shirley
Larry Frank	Philippe Cogneau
Greg Geisler	David Perkins
James Dow	Stuart Heckman
HanNa Lim	Jian Zhou
Sherman Hanna	David Hulse
Victoria Bryant	Jeremy Burke
Jeremy Clark	Cathy Faulcon Bowen
Hem C. Basnet	Yan Liu
George Korniotis	Diego Escobari
Harin Desilva	Juan Gallardo
Hal Hershfield	Jason P. Berkowitz
David Nanigian	Sharon Danes

Please consider submission to the Financial Services Review and rely on the style information provided to ease readability and streamline the review process. The Journal welcomes articles over the range of areas that comprise personal financial planning. While FSR articles are certainly diverse in terms of topic, data, and method, they are focused in terms of motivation. FSR exists to produce research that addresses issues that matter to individuals. I remain committed to the goal of making Financial Services Review the best academic journal in individual financial management and personal financial planning.

Best regards,
Stuart Michelson
Editor *Financial Services Review*

FINANCIAL SERVICES REVIEW

Financial Services Review 28 (2020) 1-16

The decrease in life insurance ownership: Implications for financial planning

Kyoung Tae Kim[a], Travis P. Mountain[b], Sherman D. Hanna[c,*], Namhoon Kim[d]

[a]*Department of Consumer Sciences, 312 Adams Hall, Box 870158, University of Alabama, Tuscaloosa, AL 35487, USA*
[b]*Department of Agricultural and Applied Economics, Virginia Tech University, 250 Drillfield Drive, Blacksburg, VA 24061, USA*
[c]*Department of Human Sciences, Ohio State University, 1787 Neil Avenue, Columbus, OH 43210, USA*
[d]*Korea Rural Economic Institute, 601 Bitgaram-ro, Naju-si. Jeollanam-do, 58321, South Korea*

Abstract

Based on our analyses of Survey of Consumer Finances datasets, the proportion of households owning a life insurance policy decreased from 72% in 1992 to 60% in 2016. We estimated logistic regressions on the likelihood of ownership of any, term, and cash value life insurance. We conclude that changes in household characteristics accounted for the decrease in term life insurance ownership, but not for the decreases in any and in cash value life insurance ownership. We also found a positive association between use of a financial planner and life insurance ownership. We discuss implications for financial planning.

JEL classification: D12; D14; G22

Keywords: Life insurance; Declining insurance demand; Survey of Consumer Finances

1. Introduction

Life insurance is an important component of risk management and insurance planning, and is also important in other financial planning topics, including employee benefits, income tax planning, and estate planning. "As late as 1960, life insurance was the

* Corresponding author. Tel.: +1-614-292-4584; fax: +1-614-292-4339.
E-mail address: Hanna.1@osu.edu (S.D. Hanna)

substance of financial planning. . . " (Brandon and Welch, 2009, p. 2). However, the importance of life insurance seems to be decreasing in terms of ownership rates, and the relative number of life insurance agents (Scism, 2016). Also, the Certified Financial Planner Board (2015) list of principal knowledge topics weight for risk management and insurance for the CFP Exam, has decreased from 14% (Hanna, et al., 2011) to 12% (Certified Financial Planner Board, 2015). A *Wall Street Journal* article (Scism, 2016) suggested that because of consolidation and decreasing ownership rates, the life insurance agent may be going the way of the dinosaur. Cordell, Finke, and Lemoine (2007) noted the long-term trend of decreasing ownership of cash value life insurance between 1995 and 2004, and Retzloff (2010) reported that life insurance ownership in the United States was at a 50 year low. However, life insurance is still a vital safety net for U.S. households. The premature death of a wage-earner is one of the more serious financial risks a household faces. Income replacement and burial expenses are the top two reasons households own life insurance (Durham, 2015).

The main purpose of our research was to ascertain factors related to the decline of life insurance ownership rates of U.S. households. We analyzed a combination of 1992 to 2016 Survey of Consumer Finances (SCF) datasets, and found that ownership of life insurance decreased until 2013 and then stayed about the same in 2016. We conducted logistic regression analyses of the pooled dataset to analyze the extent to which changes in the composition of U.S. households might have contributed to the decreases in any, term, and cash value life insurance. Further, we conducted additional analyses to investigate the role of financial planner use on life insurance ownership. Our study contributes to some implications for financial planners by providing insights into life insurance ownership trends.

2. Literature review

2.1. Overview of life insurance

Without life insurance, most families would need to reduce their current standard of living in the event of the death of a spouse or partner (Auerbach and Kotlikoff, 1991; Bernheim, Carman, Gokhale, and Kotlikoff, 2003; Bernheim, Forni, Gokhale, and Kotlikoff, 2003). Cash value and term life insurance are the two main categories of life insurance. In addition to the insurance aspect of a cash value life policy, it also provides a savings component and is intended to last the insured's "whole" life. Term policies are pure insurance that are associated with a set number of years ranging from one to 40 with 20 years being a common term contract length. If the insured dies during this contract period, the beneficiary receives the face amount of the policy. If one dies after the contract expires, no benefits are paid to the beneficiary. Employer-provided group life insurance is one of the most common types of employer benefits, and in 2017, 60% of employers provided it, with 73% of eligible employees participating in group life insurance (Greenwald and Fronstin, 2019).

2.2. Previous studies on life insurance ownership

Mulholland, Finke, and Huston (2016) examined the declining trend in cash value life insurance ownership rates using the 1992 to 2010 waves of the SCF. The proportion of cash value life policies relative to term life policies dramatically decreased over this period. They found a substitution effect, as those who owned term life insurance were less likely to own cash value life insurance. However, Frees and Sun (2010) found a complementary effect, as those who owned term life insurance were more likely to own cash value life insurance. Mulholland et al. (2016) and Glazer (2007) noted that in addition to the income protection features of term life insurance, cash value life insurance policies have been marketed as tax-advantaged investments. However, with the introduction of more tax-advantaged savings instruments such as Roth IRAs, along with the introduction of tax-advantaged savings plans for college costs, many households now have attractive alternatives to cash value life insurance for important financial goals. Increases in federal income tax marginal rates could potentially decrease the demand for cash value life insurance. Mulholland et al. (2016) noted that cash value life insurance has had an important role in estate planning tools, especially when more households were potentially subject to the federal estate tax. The increases in the exemption amounts for the federal estate tax have generally reduced the number of households potentially subject to the tax after 2004. Mulholland et al. (2016) also noted that the cost of term life insurance has dropped substantially since the introduction of internet marketing and price comparisons, but the cost of cash value life insurance has not dropped as much. Heo, Grable, and Chatterjee (2013) examined the 2004 and 2008 National Longitudinal Survey of Youth 1979 cohort that consisted of respondents aged 43–51. Heo et al. (2013) noted that there was a net decrease of one percentage in life insurance ownership over this period.

Guillemette, Hussein, Phillips, and Martin (2015) analyzed the 1992 through 2010 SCF datasets, finding a decreasing ownership trend and examined if household size affected life insurance ownership differently for minority households. Compared with White households, as household size increases, the likelihood of life insurance ownership decreases for both Black and Hispanic households. Additionally, they found that being employed, age, education, presence of children, homeownership, and being married to be positively associated with having life insurance. Relative to 1992, all subsequent survey years were significant and negative with the exception of 1995 that was not significantly different than 1992. They also found that self-employed households were less likely to own life insurance compared with those working for an employer, and Hispanic households were less likely to own life insurance than otherwise comparable White households, while Black households were more likely to own life insurance than White households.

Gutter and Hatcher (2008) examined the demand for life insurance using the 2004 SCF, with an objective of determining factors related to White and Black household life insurance ownership differences. Gutter and Hatcher (2008) found age to be positively correlated to life insurance ownership. Homeownership was positively related to life insurance ownership while the likelihood of life insurance ownership for those with children was not different from those without children. Mountain (2015), using the 2013 SCF, examined life insurance ownership for working coupled households aged 30–64. Mountain (2015) estimated that the

median proportion of insurable human wealth insured with life insurance was only 28%. Age and education were positively associated with life insurance ownership. Households with a Black respondent were more likely, and those with Hispanic and Asian/other respondents were less likely, to own life insurance than households with a White respondent. Married couples were more likely to own life insurance than partnered couples.

Liebenberg, Carson, and Dumm (2012) explored the determinants of life insurance demand for both term and whole life policies in a dynamic analysis where the authors used 1983–1989 SCF panel datasets in a Cragg model. They found that new parents were more likely to own life insurance than those who were not new parents, and households who had started a new job were more likely to increase their holdings of term insurance coverage than those who had not. Bernheim, Carman et al. (2003) examined financial vulnerability and life insurance coverage at all stages of the life cycle. The authors proposed two explanations on why a household with financial vulnerabilities might not be protected by life insurance: (1) young households purchased long term life insurance contracts but failed to update them as life events change; (2) actual needs had no effect on purchase.

Finke, Huston, and Waller (2009) used data from the 2004 SCF and compared the difference of life insurance purchases between people who were advised by either financial planners or brokers (dealers) to those who were not. They found that households that used financial planners and that used broker/dealers shared similar demographic characteristics. More important, the estimation results showed that people who relied primarily on financial planners were more likely to purchase adequate life insurance holdings, while the use of broker/dealers had no impact on levels of life insurance. Finally, households who were wealthier, self-employed, and home-owning tended to purchase more adequate life insurance holdings. Baek and DeVaney (2005) developed a model of term and whole life insurance ownership and amount of coverage, and tested the effects of human capital, bequest motives, and risk variables, as well as other household characteristics. They found that ownership of term life insurance decreased with age, was higher for middle income households than for low income households, and higher for homeowners than for renters, but education was not significantly related to term life insurance ownership. They also found that ownership of cash value life insurance was highest for those over the age of 65, and there was also a positive relationship with age controlling for other characteristics including life expectancy. Cash value life insurance ownership was also negatively related to having excellent health, positively related to the level of liquid assets, and those in the highest income tax bracket were more likely to own than those in the lowest bracket. Ownership was not significantly related to attitudes about leaving a bequest.

Using Finke et al. (2009) as framework, Scott and Gilliam (2014) focused on baby boomer life insurance adequacy before and after the 2008 financial crisis. The authors used 2004 and 2010 SCF datasets to represent before and after the 2008 financial crisis. They found that the use of a financial planner was higher among household with adequate life insurance. In addition, having a financial planner acted as a positive predictor of life insurance adequacy in 2004, while it had no impact on life insurance adequacy in 2010, after the 2008 financial crisis. Mountain (2015) found that households who consulted a financial planner or broker were more likely to own life insurance and also, given ownership, to have higher face value

amounts, however, he did not find the financial planner or broker variable to be significant when exploring the proportion of insurable human wealth that was insured by life insurance.

2.3. Overview of past research on life insurance

We reviewed selected research on factors related to life insurance ownership and adequacy. While previous researchers have found a number of household characteristics to be related to life insurance ownership, none of the studies have had a focus on identification of factors related to changes in life insurance ownership over many years. We estimated models of life insurance ownership with a combined dataset of 1992 through 2016 SCF datasets; thus, using more recent data than previous research, some of which included datasets only through 2010. By estimating a time trend controlling for household characteristics, we investigated the question of whether the decrease in life insurance ownership has been primarily because of changes in household characteristics, or to some other factors, such as the increasing availability of alternative tax advantaged investment options such as 401k accounts (Mulholland et al., 2016).

3. Methods

3.1. Data and sample selection

We used a pooled dataset from nine waves of the Survey of Consumer Finances (SCF), from 1992 to 2016. The Federal Reserve Board (FRB) has released the SCF cross-sectional dataset triennially since 1983. Hanna, Kim, and Lindamood (2018) provide detailed information about using SCF datasets. For more information about specific variables, see Board of Governors of the Federal Reserve System (2014, 2017). The total sample size of the pooled dataset is 44,634. There were 883 cases where life insurance ownership was missing (shadow variables with values over 90, see Hanna et al., 2018), and those cases were excluded from our analyses. The final analytic sample was 43,751.

3.2. Measurement of variables

3.2.1. Dependent variable

The SCF has life insurance ownership variables for both term life insurance and cash value life insurance. For our empirical analysis as well as the life insurance ownership trend, we use three separate binary dependent variables: whether the respondent or any family member has, term, cash value, or any life insurance policy (i.e., term or cash value). The actual questions are presented in Appendix 1.

3.2.2. Independent variables

Following the existing literature on life insurance ownership, the set of independent variables included survey year, age of the household head, age squared, marital status (married, single male, single female, and partnered), education of the household head (less

than high school, high school, some college, bachelor degree, and postbachelor degree), race/ethnicity (White, Black, Hispanic, and Asian/others), presence of a child under 18, current income compared with normal income (low, normal, and high), log of household income, home ownership, employment status of the household head (salary worker, self-employment, retired, and not working) and health status of a household head (excellent, good, fair, and poor). In the SCF, race/ethnicity is of the respondent (Lindamood, Hanna, and Bi, 2007), but for convenience, we refer to the race/ethnicity of the household, for example, White households.

The SCF has had a question about whether a household reported using a financial planner for information on (1) saving and investment decisions and/or (2) borrowing and credit decisions since the 1998 SCF. We used an indicator for comprehensive financial planner use, defined as a household using a financial planner for saving/investment and/or borrowing/credit issues (Elmerick, Montalto, and Fox, 2002).

3.2.3. Analysis

Given the binary dependent variables of life insurance ownership, logistic regression models are utilized to analyze factors related to the ownership of any life insurance, term life insurance, and cash value life insurance. Because of the nature of the data and our analysis, we are not able to attribute specific causal relationships for life insurance ownership. For descriptive purposes, means tests are also employed. We used the Repeated-Imputation Inference (RII) method with all of the five implicates in each SCF dataset, which provides an estimate of variances more closely representing the true variances than estimates obtained by only one implicate (Hanna et al., 2018; Lindamood et al., 2007). In order to provide conservative hypothesis tests (Shin and Hanna, 2017), we did not weight the multivariate analyses.

4. Results

4.1. Descriptive results

Fig. 1 displays the historical downward trend in life insurance ownership from the combined 1992–2016 SCF datasets. The proportion of households owning some type of life insurance peaked at 72% in 1992 and reached a low of 60% in 2013, with the 2016 rate about the same as 2013. The proportion of households owning a term life insurance policy had a generally downward trend, starting at 54% in 1992, and decreasing to 48% by 2013, and the rate remained about the same in 2016. The proportion of households owning a cash value life insurance policy decreased from 36% in 1992 to 20% in 2013, and the 2016 rate was about the same as 2013. Table 1 presents the logistic regression results of the association between survey year and the likelihood of three different types of life insurance ownership, not controlling for other household characteristics. There were significant negative relationships between the survey year and the likelihood of ownership of any life insurance, term life insurance, and cash value life insurance.

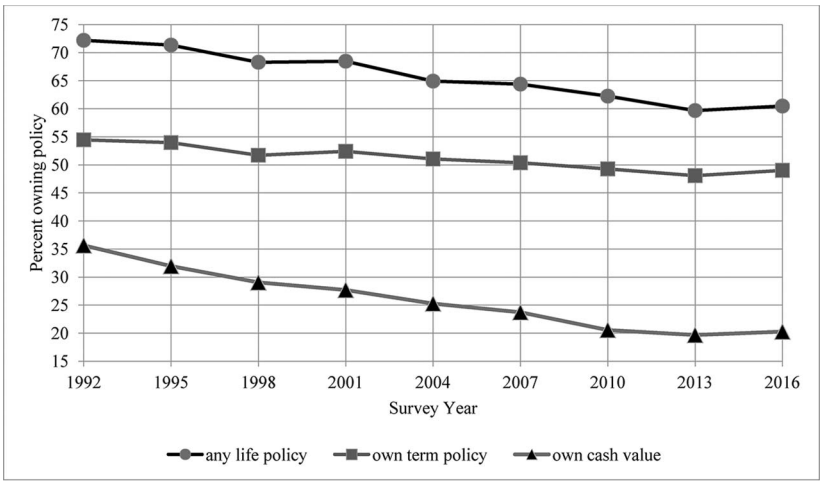

Fig. 1. Trend in life insurance ownership rates, any type, term, and cash value, 1992–2016 SCF weighted results. $N = 43{,}751$.

4.2. RII means test

The characteristics of our analytic sample are presented in Table 2. Most respondents were White, homeowners, did not have a child under 18, had current income about the same as normal, salary workers, completed education beyond the high school level and had good or excellent health status. Results of means tests are also included to show descriptive patterns of any, term, and cash value life insurance ownership by selected household characteristics. In the discussion below, we report patterns for ownership of any life insurance. The life insurance ownership rate was 46% for households with a respondent under 30, 71–72% for those with a respondent 40 to 49 and 50 to 59, and 63% for those with a respondent 70 and over. White households had a higher rate of life insurance ownership (69%) than each of the other racial/ethnic groups, and Hispanics had the lowest rate, 37%. Homeowners had a higher ownership (75%) than renters (46%).

The proportion of life insurance ownership was higher for married couples (77%) than for other groups. Households with a dependent child under age 18 had a higher life insurance ownership rate (68%) than those without a child (64%). Households with unusually high

Table 1 Logistic regressions on ownership of any, term, and cash value life insurance by survey year only, 1992–2016 SCF

	Any life insurance		Term life insurance		Cash value life insurance	
	Coefficient	Standard error	Coefficient	Standard error	Coefficient	Standard error
Survey year	−0.0270***	0.0030	−0.0072***	0.0028	−0.0394***	0.0030
Intercept	54.9622	6.0089	14.3913	5.5264	78.1899	6.0701
Mean concordance	50.1%		45.8%		53.1%	

Note: $*p < 0.05$, $**p < 0.01$, $***p < 0.001$. Unweighted analysis with RII technique. $N = 43{,}751$.

Table 2 Proportion of ownership of any, term, and cash value life, by household characteristics, 1992–2016 SCF

Variable	Category	Distribution (%)	Any life insurance		Term life insurance		Cash value life insurance	
			Rate of insurance ownership (%)	Significance level	Rate of insurance ownership (%)	Significance level	Rate of insurance ownership (%)	Significance level
Age of head	**Less than 30**	12.8	45.5	Reference	37.7	Reference	12.7	Reference
	Between 30 and 39	19.0	63.7	<0.001	54.4	<0.001	20.0	<0.001
	Between 40 and 49	20.5	71.1	<0.001	60.1	<0.001	23.6	<0.001
	Between 50 and 59	18.0	71.5	<0.001	58.4	<0.001	28.0	<0.001
	Between 60 and 69	14.0	70.2	<0.001	49.6	<0.001	33.5	<0.001
	70 and over	15.7	62.9	<0.001	37.3	<0.001	32.7	<0.001
Race/ethnicity	**White**	73.2	68.7	Reference	53.4	Reference	27.6	Reference
	Black	13.5	66.5	<0.001	50.5	<0.001	24.7	<0.001
	Hispanic	9.2	36.9	<0.001	31.8	<0.001	8.6	<0.001
	Asian/other	4.1	59.0	<0.001	47.8	<0.001	18.8	<0.001
Homeowner	**No**	33.7	45.5	Reference	36.6	Reference	13.7	Reference
	Yes	66.3	75.0	<0.001	58.0	<0.001	30.9	<0.001
Marital status	**Married**	50.3	77.1	Reference	61.4	Reference	31.6	Reference
	Single male	14.9	51.7	<0.001	38.9	<0.001	18.2	<0.001
	Single female	27.3	53.7	<0.001	39.9	<0.001	19.5	<0.001
	Partner	7.5	52.2	<0.001	43.2	<0.001	16.2	<0.001
The presence of child under 18	**No**	66.2	63.8	Reference	47.4	Reference	26.1	Reference
	Yes	33.8	67.6	<0.001	57.5	<0.001	23.3	<0.001
Current income relative to normal	**Normal**	73.1	67.1	Reference	52.3	Reference	26.1	Reference
	Income higher	8.5	71.4	<0.001	56.9	<0.001	28.4	<0.001
	Income lower	18.4	54.0	<0.001	42.2	<0.001	19.8	<0.001
Employment status of head	**Salary worker**	58.0	70.7	Reference	59.4	Reference	23.6	Reference
	Self-employment	10.9	65.4	<0.001	48.8	<0.001	31.5	<0.001
	Not working	5.7	33.8	<0.001	25.5	<0.001	12.5	<0.001
	Retired	25.3	59.0	<0.001	37.6	<0.001	28.9	<0.001
Education of household head	**less than high school**	14.6	45.6	Reference	33.5	Reference	16.3	Reference
	High school degree	30.7	63.1	<0.001	47.7	<0.001	24.6	<0.001
	Some college	23.0	66.6	<0.001	52.2	<0.001	25.4	<0.001
	Bachelor degree	18.6	74.7	<0.001	61.0	<0.001	28.7	<0.001
	Post-bachelor degree	11.5	76.7	<0.001	61.6	<0.001	32.2	<0.001
Health status of head	**Excellent health**	27.1	70.2	Reference	56.6	Reference	26.8	Reference
	Good health	47.3	66.9	<0.001	52.9	<0.001	25.6	<0.001
	Fair health	19.4	58.4	<0.001	42.8	<0.001	23.3	<0.001
	Poor health	6.2	49.7	<0.001	34.9	<0.001	20.1	<0.001

Note: Author analyses of combined 1992–2016 SCF datasets, $N = 43{,}751$. RII technique is used for significance tests. The reference category used in the means test is indicated in bold face. Significance test is for mean difference from reference category for each variable.

current income were more likely to own life insurance (71%) than those with lower current income (54%) and those with current income about the same as normal (67%). Households with the head working for a salary had a higher rate of life insurance ownership (71%) than those in other types of employment status, and households with the head not working had only a 34% life insurance ownership rate. The proportion of life insurance ownership was highest for households with a head having a postbachelor degree (77%), compared with 75% for households with a bachelor degree, 67% for households with some college, 63% for those with a high school degree, and only 46% for those less than high school degree. Households with a head with excellent health had the highest proportion of life insurance, 70%, while those with poor health had the lowest proportion, 50%.

Table 3 Logistic regression analysis of likelihood of ownership of any, term, and cash value life, 1992–2016 SCF

	Any life insurance		Term life insurance		Cash value life insurance	
	Coefficient	Standard error	Coefficient	Standard error	Coefficient	Standard error
Survey year	−0.0277***	0.0034	−0.0039	0.0030	−0.0422***	0.0033
Age of head	0.0776***	0.0093	0.0838***	0.0092	0.0510***	0.0108
Age squared (/10000)	−6.3726***	0.8658	−8.6443***	0.8716	−2.6687**	0.9693
Race/ethnicity (reference: White)						
Black	0.5309***	0.0862	0.2519**	0.0793	0.3662***	0.0903
Hispanic	−0.9927***	0.0972	−0.7450***	0.0957	−0.8456***	0.1414
Asian/other	−0.3491**	0.1244	−0.2491*	0.1155	−0.2701*	0.1330
Homeowner	0.7555***	0.0636	0.5241***	0.0614	0.5487***	0.0738
Marital status (reference: married)						
Single male	−0.7504***	0.0793	−0.5418***	0.0754	−0.4549***	0.0864
Single female	−0.7767***	0.0689	−0.5186***	0.0651	−0.5665***	0.0753
Partnered	−0.6230***	0.1021	−0.4268***	0.0970	−0.3280**	0.1205
The presence of child under 18	0.1236	0.0657	0.1713**	0.0582	0.0739	0.0646
Current income relative to normal (reference: Normal)						
Income Higher	0.0294	0.0898	−0.0272	0.0762	0.1150	0.0799
Income lower	−0.3024***	0.0671	−0.2439***	0.0628	−0.1013	0.0724
Log of household income	0.0547***	0.0138	0.0218	0.0129	0.0451**	0.0147
Employment status of head (reference: Salary worker)						
Self-employment	−0.5130***	0.0736	−0.6081***	0.0624	0.2854***	0.0646
Not working	−1.1669***	0.1182	−1.0809***	0.1190	−0.3563*	0.1546
Retired	−0.7338***	0.0883	−0.6588***	0.0793	−0.0608	0.0854
Education of head (reference: Less than high school)						
High school degree	0.3858***	0.0842	0.1921*	0.0828	0.4210***	0.1008
Some college	0.5439***	0.0901	0.3173***	0.0870	0.4984***	0.1050
Bachelor degree	0.7273***	0.0951	0.4433***	0.0895	0.6248***	0.1055
Post-bachelor degree	0.5862***	0.1015	0.4375***	0.0942	0.5248***	0.1092
Health status of head (reference: Excellent health)						
Good health	0.0215	0.0615	0.0278	0.0537	0.0007	0.0574
Fair health	−0.0697	0.0813	−0.0677	0.0747	−0.0392	0.0831
Poor health	−0.2140	0.1256	−0.0945	0.1239	−0.2482	0.1422
Intercept	53.3805	6.7225	5.6963	6.0027	80.5698	6.5469
Mean concordance	75.6%		70.7%		72.0%	

Note: $*p < 0.05$, $**p < 0.01$, $***p < 0.001$. Unweighted analysis with RII technique. $N = 43{,}751$.

4.3. Multivariate results

Table 3 shows the logistic regression results of the likelihood of three different types of life insurance ownership from the 1992–2016 SCF dataset, controlling for household characteristics. There was a negative relationship between the survey year and the likelihood of life insurance ownership for any life insurance and cash value life insurance, but the relationship was not significant for term life insurance ownership. The magnitudes of the coefficients of the time trend for ownership of any life insurance, and also for ownership of cash value life insurance were very similar between the results in Table 3 and the results in Table 1, suggesting that the negative trends in ownership and in cash value life insurance were not because of changes in household characteristics. However, the effect of the time trend on term life insurance ownership was not significant in Table 3, and the magnitude of

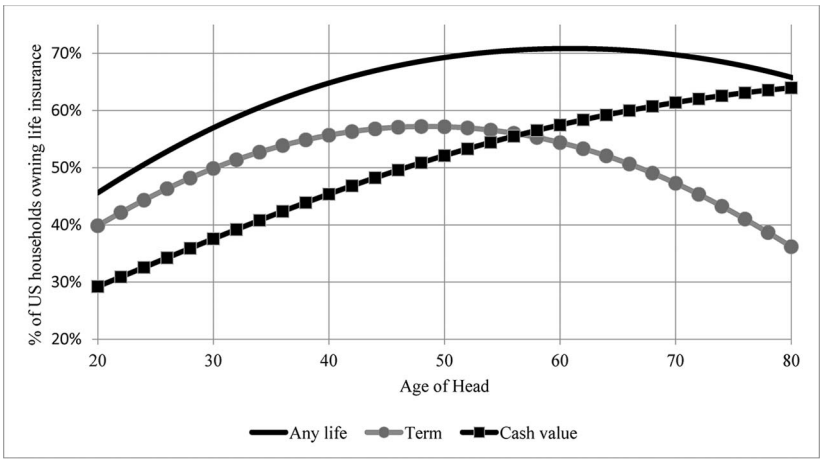

Fig. 2. Effect of age on the proportion of households owning any, term, and cash value life insurance, assuming other household characteristics have 2016 mean levels. Estimates based on logistic regressions in Table 3.

the effect was much less than the magnitude in Table 1, suggesting that the decrease in term life insurance might have been because of changes in household characteristics from 1992 to 2016.

Both age and age squared were significantly related to the likelihood of insurance ownership in each regression. The combined effect of the age variables implies that at the 2016 mean values of other independent variables, the likelihood of owning any life insurance increased with age until age 61, then decreased with age; ownership of term life insurance increased with age until age 48, then decreased with age; and ownership of cash value life insurance increased with age for all ages in the sample. Fig. 2 shows the calculated relationship between age of the head and the likelihood of ownership of each type of life insurance, based on the coefficients in Table 3 and assuming the 2016 mean values of other independent variables. This pattern is plausibly related to the cost of life insurance, which increases substantially with age. Households who purchased 20-year term policies in their 40s and 50s would have expiring policies in their 60s and 70s. At this point in life the cost of the insurance may outweigh the benefits, particularly if the household has relatively little debt and has adequate assets to offset the loss of income resulting from the death of a spouse or significant other. On the other hand, given estate planning goals and reduced opportunities for tax sheltered retirement accounts after age 70, the estimated increase in ownership with age of cash value ownership is plausible, since household income is assumed to remain at the overall 2016 sample mean.

For the remaining independent variables in Table 3, we discuss only the effects for ownership of any life insurance, though for many of the variables, the effects were similar for ownership of term life insurance and of cash value life insurance. For racial/ethnic patterns based on the regressions in Table 3, assuming other household characteristics had the mean levels for 2016, Black households had the highest predicted likelihood of owning life insurance, over 73%, compared with 61% for White households, 53% for Asian/other households, and only 36% for Hispanic households. There is no theoretical reason for there

to be differences in life insurance ownership between difference race/ethnic groups, if other household characteristics are equal. The reason for these differences may be related to differences in marketing practices of insurance companies. Hispanic and Asian households consist of a higher proportion of immigrant households, so it is possible that insurance companies have not effectively found a way to market to these individuals (Mountain, 2015), especially if life insurance is a product that is not well established in their country of origin. Hispanic and Asian households may also rely more heavily on their extended family network than both Black and White households, something that is not measured in the SCF.

For the relationship between marital status and life insurance ownership, assuming other household characteristics had the mean levels for 2016, married households had the highest predicted likelihood of owning life insurance, almost 70%, compared with 56% for partner households, 50% for single male households, and 49% for single female households. Surviving individuals in a coupled household are likely to be financially burdened if a spouse or partner were to die. Life insurance can help fill this burden. In general, single headed households are not financially dependent on another wage-earner and have little need for life insurance, because there would be no surviving spouse, though as Nam and Hanna (2019) discussed, there might still be a need in terms of making sure the person with custody of dependent children would have adequate resources. The difference between married and partnered coupled households may be because of the married households having a longer-term time horizon than otherwise similar partnered households. Presence of a dependent child under age 18 was not significantly related to ownership of any life insurance or of cash value life insurance, but was positively related to ownership of term life insurance. This is reasonable, because dependent children make the need for support for them paramount.

Household income was positively associated with the likelihood of owning any life or of owning cash value life insurance. For the relationship between income shocks and life insurance ownership, assuming other household characteristics had the mean levels for 2016, households with current income below normal had the lowest predicted likelihood of owning life insurance, 54%, compared with 64% for households with higher than normal income and 61% for households with normal income. This may be because of household budget constraints. Households may accurately perceive the likelihood of death in the current year as very low, and, thus, temporarily terminate a life insurance policy if household income falls below normal income. Conversely, when household income is above normal income, households may feel less of a constraint on the household budget and use this extra income on life insurance, something they otherwise were not willing to own. Current income was positively related to the likelihood of owning any life insurance and to the likelihood of owning cash value life insurance, but not to the likelihood of owning term life insurance. Given that employment status is controlled, the lack of significance is reasonable, as an employer group life insurance is something that is available for many employees without having to make a purchase decision.

For the relationship between employment status of the head and life insurance ownership, assuming other household characteristics had the mean levels for 2016, households headed

Table 4 Logistic regression analysis, the association between comprehensive financial planner use and the likelihood of life insurance ownership, 1998–2016 SCF

	Any life insurance		Term life insurance		Cash value life insurance	
	Coefficient	Standard error	Coefficient	Standard error	Coefficient	Standard error
Use of comprehensive financial planner	0.3727***	0.0425	0.2117***	0.0356	0.2989***	0.0364
Survey year	−0.0230***	0.0021	0.0169	0.0020	−0.0398***	0.0022
Intercept	44.0071***	4.2934	−5.6204	3.9183	75.8085***	4.3208
Other control variables	Yes		Yes		Yes	
Mean concordance	75.9%		71.3%		71.7%	

Note: $*p < 0.05$, $**p < 0.01$, $***p < 0.001$. Unweighted analysis with RII technique. $N = 35{,}667$. Control variables include the independent variables shown in Table 3.

by an employee had the highest predicted likelihood of owning life insurance, 68%, presumably partly because of the availability of group life insurance from many employers. Group insurance may be provided for free or at reduced rate as an employer benefit. Additionally, group life insurance removes the barrier of having to have a medical exam to be insured, something that is common in the individual insurance market. This barrier may prevent individuals from seeking life insurance, even if they would qualify for the policy upon completion of the exam. The predicted life insurance ownership rates were only 51% for households with a self-employed head, 49% for households of a retired head, and 40% for households with heads not working but not of retirement age.

For the relationship between the education of the head and life insurance ownership, assuming other household characteristics had the mean levels for 2016, households headed by somebody with a postbachelor degree had a predicted life insurance ownership rate of 65%, and those with a bachelor degree had a predicted rate of 68%, compared with 62% for those with some college but no degree, 57% for those whose highest education is a high school degree, and 44% for those with no high school degree. Higher education levels are likely related to more future oriented thinking (Yuh and Hanna, 2010), which would be consistent with the life insurance ownership patterns shown in Fig. 10.

For the relationship between homeownership status and life insurance ownership, assuming other household characteristics had the mean levels for 2016, homeowners had a predicted ownership rate of 67%, compared with a predicted rate of 47% for renters. Homeowners with a mortgage may need to have life insurance to make sure the loan is paid off if a primary earner dies. Homeownership may also be a signal of financial stability. Typically, homeowners are required to make a substantial down payment to purchase a home, which, for most households requires substantial financial planning. Therefore, homeowners may be more likely to both afford and plan for life insurance.

4.4. The association between use of financial planner and life insurance ownership

Table 4 shows the logistic regression results of the association between the use of comprehensive financial planner and the likelihood of different types of life insurance

ownership, controlling for the other independent variables in Table 3. The SCF did not have a specific code for use of financial planners before 1998, so we used a pooled dataset from the 1998–2016 SCF for this additional analysis. The utilization of a comprehensive financial planner was positively associated with ownership of any, term, and cash value life insurance policies. Households who used a comprehensive financial planner had 45.2% higher odds of owning any life insurance than those without using a financial planner. Similarly, the use of comprehensive financial planner increased the odds of owning term life and whole life insurance policy by 24% and 35%, respectively. Controlling for use of a financial planner did not substantially change the effect of the survey year compared with the results in Table 3, as the survey year had a negative effect for any life insurance and for cash value life insurance, but the effect was not significantly different from zero for term life insurance.

5. Discussion and implication

The effect of the time trend in the logistic regression means that even if the characteristics of households in the United States had not changed between 1992 and 2016, there would have been a substantial decrease in ownership of any life insurance and of cash value life insurance during that period. However, based on the logistic regression in Table 3, ownership of term life insurance might not have decreased if household characteristics had not changed. The trends for some household characteristics related to lower likelihood of life insurance ownership (Hispanic and Asian/other households, unmarried households) might have contributed to lower ownership during the 1992 to 2016 period, but the trends for education and age might have contributed to increased ownership. Overall, changes in household characteristics during the periods played a relatively small role in the decreases in ownership of any and of cash value life insurance, partly because most household characteristics changed slowly. For instance, the proportion of households with married couples decreased from 53.9% in 1992 to 46.8% in 2016, and the proportion of household heads who were employees (vs. self-employed or retired or not working) changed from 55% in 1992 to 56% in 2016.

Insuring human wealth is an appropriate life insurance objective (Chen, Ibbotson, Milevsky, and Zhu, 2006) but even at the current ownership rates, the amount of life insurance is not enough to replace lost human wealth for many households (Mountain, 2015). While the cash-value policy decline may reflect rational decision-making because of changes in tax laws and the increase in tax advantaged investment options, these patterns could change with changes in the federal income tax and with any substantial changes in the federal estate tax.

Durham (2015) found that households often overestimate the cost of life insurance and do so by drastic amounts. When households were asked to provide what they thought a $250,000 20-year term policy would cost, one-quarter overestimated the cost by 625% and one half overestimated the cost by 250%. Meanwhile, life insurance advice from a financial planner is something that families place a low value on (Warschauer and Sciglimpaglia, 2012), making life insurance planning something a family is more likely

to do on its own. While we found higher likelihood of life insurance ownership for households who have a financial planner, we cannot be certain if the financial planner's advice led to life insurance ownership or if households who have life insurance were more likely to seek advice from a financial planner. Future research might use methods to correct for the selection effect, as was done by Kim, Pak, Shin, and Hanna (2018) for analyses of the effect of financial planner use on holding a retirement savings goal.

It is of little surprise that life insurance ownership trends are decreasing when households place little importance on life insurance planning while also dramatically overestimating its cost. If financial planners were to view term life insurance as a loss leader product, or a breakeven product, instead of a profit generator, both the financial planner and client might be better off. As a gateway product, term insurance can help establish trust and rapport with clients. Planners could start by asking the client how much they think a $100,000 20-year term life policy would cost them, and see how that compares with reality. Once the client has the peace of mind that his or family will be financially taken care of upon an untimely death, relatively inexpensively, transition to other planning areas can ensue. This approach would likely benefit financial planners who are following or not following a fiduciary standard, recommendations of term life insurance without direct financial benefit may diminish apprehension some clients might otherwise feel may have when presented with the advice to purchasing life insurance. For clients with more complicated financial needs where cash-value life insurance is appropriate, it may be a product that is brought up later, rather earlier, in the financial plan.

If the ownership rates continue to decline, more and more families will face financial hardship and economic distress if a spouse or partner dies prematurely. In terms of the potential benefits of financial planning advice based on normative economic models, risk management is an important component of the value of advice (Hanna and Lindamood, 2010), and for many families, life insurance still should be a salient component of a financial plan. This is surprising as there are several studies have found that a substantial proportion of households still remain underinsured. Financial planners and educators can better inform their clients and the population as to the merits of life insurance. Additionally, alternative products, adjustments to the current life insurance products, or additional saving will need to fill the void of the disappearing life insurance.

Given our conclusion that the decreases in ownership of any life insurance and in ownership of cash value life insurance were not because of changes in household characteristics between 1992 and 2016, it is plausible that future changes in federal income tax and estate tax policies could lead to future increases in life insurance ownership. For instance, if a proposal by Senator Bernie Sanders to decrease the estate tax exemption from $11.4 million in 2019 to $3.5 million (2009 level) were to be implemented, many wealthy households would likely make substantial changes in their estate planning (Shenkman, 2019), and cash value life insurance might provide one type of strategy. Further, proposed increases in federal income tax rates might also make cash value life insurance more attractive. Even if no near-term tax legislation is passed, most individual income tax changes in the Tax Cuts and Jobs Act of 2017 are set to sunset January 1, 2026. The increases in income tax rates might nudge households to reconsider cash-value life insurance policies.

Appendix 1 Life insurance variables in the SCF

Variable	Description
X4001	Do you (or anyone in your family living here) have any life insurance? Please include individual and group policies, but not accident insurance. 1. *YES 5. *NO
X4002	The two major types of life insurance are term and cash-value policies. Term policies pay a benefit if the insured person dies, but otherwise have no value. They are often provided through an employer or union, but may also be bought by individuals. Cash-value policies also pay a death benefit, but differ in that they build up a value as premiums are paid. Are any of your (family's) policies term insurance? 1. *YES 5. *NO 0. Inap. (no life insurance: X4001^ = 1)
X4004	Do you have any policies that build up a cash value or that you can borrow on? These are sometimes called 'whole life', "straight life," or "universal life" policies. 1. *YES 5. *NO 0. Inap. (no life insurance: X4001^ = 1)

References

Auerbach, A. J., & Kotlikoff, L. (1991). *Life Insurance Inadequacy-Evidence From a Sample of Older Widows* (NBER Working Papers 3765). National Bureau of Economic Research.

Baek, E., & DeVaney, S. A. (2005). Human capital, bequest motives, risk, and the purchase of life insurance. *Journal of Personal Finance, 4,* 62–84.

Bernheim, B. D., Carman, K. G., Gokhale, J., & Kotlikoff, L. J. (2003). Are life insurance holdings related to financial vulnerabilities? *Economic Inquiry, 41,* 531–554.

Bernheim, B. D., Forni, L., Gokhale, J., & Kotlikoff, L. J. (2003). The mismatch between life insurance holdings and financial vulnerabilities: Evidence from the Health and Retirement Study. *American Economic Review, 93,* 354–365.

Board of Governors of the Federal Reserve System. (2014). *Codebook for the 2013 Survey of Consumer Finances*. (available at http://www.federalreserve.gov/econresdata/scf/files/codebk2013.txt)

Board of Governors of the Federal Reserve System. (2017). *Codebook for the 2016 Survey of Consumer Finances*. (available at https://www.federalreserve.gov/econres/files/codebk2016.txt)

Brandon, Jr., E. D., & Welch, H. O. (2009). *The History of Financial Planning: The Transformation of Financial Services.* Hoboken, NJ: John Wiley & Sons.

Certified Financial Planner Board. (2015). *2015 Principal Knowledge Topics*. (available at https://www.cfp.net/become-a-cfp-professional/2015-job-task-analysis/2015-principal-knowledge-topics)

Chen, P., Ibbotson, R. G., Milevsky, M. A., & Zhu, K. X. (2006). Human capital, asset allocation, and life insurance. *Financial Analysts Journal, 62,* 97–109.

Cordell, D. M., Finke, M., & Lemoine, C. (2007). The death of life? The demand for cash-value life insurance has slipped in recent years, say the authors, much of it due to bad publicity. *Journal of Financial Planning, 20,* 42–44.

Durham, A. (2015). *2015 Insurance Barometer Study. Life Happens*. (available at https://www.lifehappens.org/industry-resources/2015barometer/)

Elmerick, S. A., Montalto, C. P., & Fox, J. J. (2002). Use of financial planners by U.S. households. *Financial Services Review, 11,* 217–231.

Federal Reserve Bank. (2016). *Codebook for 2016 Survey of Consumer Finances*. Washington, DC: Federal Reserve Bank.

Finke, M. S., Huston, S. J., & Waller, W. (2009). *Do contracts impact comprehensive financial advice?* (available at http://dx.doi.org/10.2139/ssrn.1429807)

Frees, E. W., & Sun, Y. (2010). Household life insurance demand: A multivariate two-part model. *North American Actuarial Journal, 14,* 338–354.

Glazer, D. A. (2007). The effect of commission versus non-commission benefits on customer value: The case of life insurance policy performance. *Financial Services Review, 16,* 135–153.

Greenwald, L., & Fronstin, P. (2019). *The state of employee benefits: Findings from the 2018 Health and Workplace Benefits Survey.* Employee Benefit Research Institute No. 470. (available at https://www.ebri.org/docs/default-source/ebri-issue-brief/ebri_ib_470_wbs2-10jan19.pdf?sfvrsn=c5db3e2f_10)

Guillemette, M., Hussein, M., Phillips, G. M., & Martin, T. (2014). Minority household size and the life insurance purchase decision. *Financial Services Review, 24,* 37–51.

Gutter, M. S., & Hatcher, C. B. (2008). Racial differences in the demand for life insurance. *Journal of Risk and Insurance, 75,* 677–689.

Hanna, S. D., & Lindamood, S. (2010). Quantifying the economic benefits of personal financial planning. *Financial Services Review, 19,* 111–127.

Hanna, S. D., Ji, H., Lee, J., Son, J., Letkiewicz, J., Lim, H., & Zhang, L. (2011). Content analysis of Financial Services Review. *Financial Services Review, 20* (3), 237–251.

Hanna, S. D., Kim, K. T., & Lindamood, S. (2018). Behind the numbers: Understanding the Survey of Consumer Finances. *Journal of Financial Counseling and Planning, 29,* 410–418.

Heo, W., Grable, J. E., & Chatterjee, S. (2013). Life insurance consumption as a function of wealth change. *Financial Services Review, 22,* 389–404.

Kim, K. T., Pak, T. K., Shin, S. H., & Hanna, S. D. (2018). The relationship between financial planner use and holding a retirement saving goal: A propensity score matching analysis. *Financial Planning Review, 1, e1008.* https://doi.org/10.1002/cfp2.1008

Liebenberg, A. P., Carson, J. M., & Dumm, R. E. (2012). A dynamic analysis of the demand for life insurance. *Journal of Risk and Insurance, 79,* 619–644.

Lindamood, S., Hanna, S. D., & Bi, L. (2007). Using the Survey of Consumer Finances: Some methodological considerations and issues. *Journal of Consumer Affairs, 41,* 195–222.

Mountain, T. P. (2015). *Life Insurance and Financial Vulnerability.* Ph.D. dissertation, The Ohio State University.

Mulholland, B., Finke, M., & Huston, S. (2016). Understanding the shift in demand for cash value life insurance. *Risk Management and Insurance Review, 19,* 7–36.

Nam, Y., & Hanna, S. D. (2019). The effects of risk aversion on life insurance ownership of single-parent households. *Applied Economics Letters, 26,* 1285–1288. http://dx.doi.org/10.1080/13504851.2018.1546044

Retzloff, C. D. (2010). *Household Trends in U.S. Life Insurance Ownership*. LIMRA, report, November, 1.

Scism, L. (2016, Feb 25). Death of the insurance salesman at MetLife. *The Wall Street Journal.*

Scott, J. K., & Gilliam, J. (2014). Boomers' life insurance adequacy pre & post the 2008 financial crisis. *Financial Services Review, 23,* 287–304.

Shenkman, M. (2019, March 21). *Sanders Estate Tax Proposal: Estate Planning Steps to Take Now.* (available at https://www.forbes.com/sites/martinshenkman/2019/03/21/sanders-estate-tax-proposal-estate-planning-steps-to-take-now/#16bae6299f5e)

Shin, S. H., & Hanna, S. D. (2017). Accounting for complex sample designs in analyses of the Survey of Consumer Finances. *Journal of Consumer Affairs, 51,* 255–283.

Warschauer, T. A. (2012). The economic benefits of personal financial planning: An empirical analysis. *Financial Services Review, 21,* 195–208.

Yuh, Y., & Hanna, S. D. (2010). Which households think they save? *Journal of Consumer Affairs, 44,* 70–97.

FINANCIAL SERVICES REVIEW

Financial Services Review 28 (2020) 17-34

Are "fun" sources of windfalls destined to be spent hedonistically?

Eugene Bland[a]*, Valrie Chambers[b]

[a]*Texas A&M University – Corpus Christi, 6300 Ocean Drive, Corpus Christi, TX 78412, USA*
[b]*Stetson University, 421 N Woodland Boulevard Unit 8398, DeLand, FL 32723, USA*

Abstract

Recently, Richard Thaler was awarded a Nobel Prize for his work in developing Behavioral Economics. While much of economics assumes that people act rationally, Areily (2008), expanding on Thaler's body of work, proves that we are not only often irrational, but we are predictably irrational. When an interviewer asked Thaler how he would spend the roughly $1.1 million in prize money, he responded, "This is quite a funny question." Thaler added, "I will try to spend it as irrationally as possible." We know that affective tags for money exist but what specifically are those affective tags? More specifically still, is one of those tags for sources of income "fun," and if so, does that affect whether the money will be spent on fun? Classical economics would assume that satisfaction comes from the consumption of goods and services, that money is a medium of exchange, and that the source of that medium of exchange does not enter into the choice of the goods or services consumed. Thaler's (1999) works show that people create mental accounts, indicating that the source of the money may not be as completely irrelevant as classical economics predicted. This is important because where irrational behavior is suboptimal behavior, if we can anticipate it, we can construct environments to support better choices. We find that fun sources of income are significantly more likely to be spent on fun expenditures. However, as the amount of the windfall increases, the amount spent on fun levels off, indicating that this affect may be bounded. We were unable to find statistically significant support that more "adult" sources of income are more likely to be spent on more adult uses, but money from adult sources was significantly more likely to be invested. This is important because understanding more about affective tags and how they affect decisions to use money, we become better predictors of irrational behavior.

JEL classification: G40; D140

Keywords: Income; Income source; Behavioral finance; Consumer behavior; Mental accounting

* Corresponding author. Tel.: +1-361-825-2829; fax: +1-361-825-5609.
E-mail address: Eugene.Bland@tamucc.edu (E. Bland)

1. Introduction

Economic researchers have traditionally assumed that people's behavior is rational. Classical economics would assume that satisfaction comes from the consumption of goods and services, that money is a medium of exchange, and that the source of that medium of exchange does not enter into the choice of the goods or services consumed. Dan Ariely's (2008) work asserts however that we are predictably irrational. However, while deviations from rational behavior abound, behavioral economics is relatively new and there is much to learn about how rationality is bounded and how people make financial decisions. Frederick (2005) found that the more cognitive reflection that occurred, the less nonrational behavior occurred. Thaler's (1999) work shows that people create mental accounts, indicating that the source of the money may not be as completely irrelevant as classical economics predicted. That is, there is much to be learned about how rationality and irrationality interact.

Specifically, we know that affective tags for money exist (Bradford, 2008; Henderson and Peterson, 1992; Levav and McGraw, 2009; Winkelmann et al., 2011,) but we do not know what specifically those affective tags are, nor do we know the strength of those tags. More specifically still, one of those affective tags for sources of income may be "fun," and if so, that may affect how the money will be spent. There may be a direct connection between the affective tag on income and the affective tag on the disposition of that income. This connection has yet to be directly studied. By understanding more about what affective tags are and how affective tags on income affect decisions to use money, we become better predictors of people's economic behavior.

Such understandings in aggregate can help lead people to create systems whereby they make better financial choices, which in turn reduces their financial stress and increases their wealth and quality of life. This research is important as part of a broader field of research because where irrational behavior is suboptimal behavior, if we can anticipate it, we can construct environments to support better choices. Alternatively, by understanding where irrational behavior may occur, policymakers may be able to provide alternative choices that may result in better outcomes. This research is important individually because as individuals better understand their own behavior, they can reflect upon it and adjust their behavior to better achieve their goals. For example, if a person knows that she generally receive a tax refund and she is prone to use such a windfall for adult purposes, she can incorporate a savings plan directly attached to that refund that will build her wealth more quickly than she currently does.

2. Literature review

According to Thaler's (1999) mental accounting theory, people create different mental accounts like long-term savings and have different marginal propensities to consume from each account. Academic literature supports mental accounting theory from a regular income flow or from an irregular, lump-sum, windfall (Adamopoulou and Zizza, 2017; Johnson et al., 2006; O'Curry, 1999; Souleles, 2002), and supports the periodic reconciliation of

people's mental accounts for income and expense (Camerer et al., 1997; Heath and Soll, 1996; Read et al., 1999; Rizzo and Zeckhauser, 2003).

Characteristics of the use of mental accounts have been studied by Karlsson et al. (1999), who reported that cash spending on a durable good depended on compatible reasons for saving. Abeler and Marklein (2016) and Benjamin (2006) found that math aptitude affects mental budgeting, and Cheema and Soman (2006) and Wertenbroch (2001), concluded that mental budgeting was a matter of self-control. Arkes et al. (1994) found that a greater percentage of a small windfall was spent than that from the same amount of anticipated ordinary income, indicating that foreknowledge of income is a factor in saving, supporting the findings of Rucker (1984) and consistent with the findings of Karlsson et al. (1999). Trump et al. (2015) found that individuals made riskier choices with a stranger's money than with a friend's money. Whether income was earned affected how compliant taxpayers were after a tax audit (Boylan, 2010), and whether earned income was a windfall or restores a status quo was found to be significant (Agarwal and Qian, 2014; Epley and Gneezy, 2007).

The size of the income can also be significant. Rucker (1984) studied the retroactive payment of a raise approved by a university, reversed by the Federal Pay Board but then reinstated by the U.S. Supreme Court and found that the size of the windfall was the most important factor for deciding how the funds were used, although the length of time that the recipient had to anticipate the income was also significant. Chambers et al. (2009) studied responses to small hypothetical tax rebates and found that at some amount over $600, materiality was significant in how the money would be used.

Karlsson et al. (1999) noted that individuals considered the future consequences of spending in their mental budgeting, indicating that the permanence of the income might be significant. Friedman's (1957) permanent income hypothesis says that people will spend money consistent with their perceived permanent income level. Similarly, Blinder (1981) posited that a permanent tax decrease would elicit more spending than a temporary tax cut. Parker (1999) found that a temporary, end-of-year reduction in the social security tax for high-income wage earners was spent when received, not averaged evenly over the fiscal year. Hsieh (2003) studied large, regular bonuses associated with the annual Alaska Permanent Fund payment, which was fully anticipated and found no spike in consumption. However, consumption by the same households was very responsive to income tax refunds, suggesting that predictable and regular payments are built into consumption decisions (Hsieh, 2003). Browning and Collado (2001) studied Spanish panel data to measure the effect of customary, predictable bonus payments and like Hsieh (2003), did not find changes in consumption.

In contrast, studies of the spending from nonrecurring, nonpermanent sources of income are fairly rare. Bodkin (1959) estimated the marginal propensity to consume from a one-time dividend paid in 1950 to World War II veterans by the National Service Life Insurance to be between 0.72 and 0.97. Kreinin (1961) analyzed the consumption of a sample of Israeli citizens receiving restitution payments from Germany in 1957 and 1958 and estimated that 35% was spent. Shapiro and Slemrod (1995) found that almost half of the respondents receiving decreased periodic tax withholding refunds in 1992 would spend them, even though the total yearly tax liability remained unchanged, resulting in a lower end-of year tax refund. However, in 2001, when a tax cut took the form of a lump-sum rebate, only about one-fourth of the respondents surveyed expected to spend the payment (Shapiro and

Slemrod, 2003). Chambers and Spencer (2008) found that the timing of payments (whether paid as a lump-sum or spread out in equal monthly installments for a year) was significant, and Sahm et al. (2012), confirmed that finding. However, whether people were in the habit of saving versus spending also mattered (Spencer and Chambers, 2012).

The framing of payments seems to matter: Baker et al. (2007) found that more money was spent from likely recurring income (dividends) than from less recurring capital gain income. Hershfield et al. (2015) found that consumers placed savings and debt into different mental accounts, making them insensitive to the significant differences between the interest rates on these accounts. Shefrin and Thaler (1988) found that more of a lump sum bonus was saved than if the same amount increased regular income, even when the bonus is fully anticipated.

(1) Is the source of the income important in mental accounting?

Windfall sources in prior literature include: inheritance (Baker and Nofsinger, 2002; Zagorsky, 2013), bonus (Henderson and Peterson, 1992; Hsieh, 2003), tax rebate (Chambers and Spencer, 2008; Meekin et al., 2015), and lottery (Winkelmann et al., 2011). Some evidence suggests that the source of one's income is important in mental accounting. Winkelmann et al. (2011) found that spending from different sources of income conferred different marginal utilities. Sources of income may be tied to uses of income. For example, Henderson and Peterson (1992) reported that individuals were more likely to spend $2,000 on a vacation if the source of the funds was a gift rather than a work bonus. Bradford (2008) found that individuals allocate gifted and inherited assets consistent with their goals in the relationship. Epley et al. (2006) found that people spent more from an income source of the same amount and timing labeled "bonus" than they did one labeled "rebate."

Milkman and Beshears (2009) found that consumers who received $10 windfalls in the form of grocery coupons spent an additional $1.59 on groceries that the consumer did not typically buy. Chambers et al. (2017) found that people given a hypothetical payment from one of five different sources would spend the funds differently, depending on the source of the money, and that less of the windfall would be saved overall from a game show payment than from a tax rebate. Similarly, Chambers et al. (2017) found that people tended not to shift away from their spending habits. The goal of this article, given that money is fungible, is to test to see whether the affective tag of the spending significantly mirrors the affective tag of the windfall source.

(2) Affective tags and mental accounting

Levav and McGraw (2009) proposed that windfalls in mental accounting may have a feeling attached to that sum of money, or "affective tag." They found that when a windfall that is negatively tagged is received, the associated negative feelings influenced respondents to consume the windfall either reluctantly or virtuously to cope with those negative feelings. O'Curry and Strahilevitz (2001) found that those receiving lotto payments spent it hedonistically. This study focuses on one of those questions: does an income source affectively tagged as "fun" result in significantly more spending on fun? Will less fun sources be used

more for less fun uses? Additionally, as a research question, how is the spending on fun bounded, if at all?

(3) Demographic factors

Several demographic factors might be significant. Chen and Volpe (2002) found that gender was significant to personal finance, but education and experience could have a significant impact on the financial literacy of both genders. Fisher et al. (2015) found that income, income uncertainty, wealth, high-risk tolerance, and savings also differed significantly by gender, as did being nonwhite and having other household members. Fisher (2010) found that certain race differences in savings were explained by the individual determinants of saving, including: receiving government assistance, feeling that credit use is bad, being turned down for credit in the past five years, or having a shorter saving horizon; race also significantly affected risk tolerance.

3. Hypothesis and research questions

This study examines whether people spend the same proportion of a distribution on fun categories when the windfall source is a fun source, like from a gameshow, as they do when the source is less fun, like from a tax rebate. This study examines the spending from tax rebates, game show winnings and bonuses, which might be a more neutral benchmark. Only windfall earnings will be explored, as literature indicates that amounts spent from windfall income is spent differently from one's regular income (Arkes et al., 1994; Karlsson et al., 1999).

How might the recipient consider these sources as similar or different? Tax systems are run by a government or its appointed agency and are largely outside the respondent's control, whereas bonuses and game show winnings are generally run by private enterprises and may have more elements of respondent's control. To what extent the money is "earned" is debatable, but bonuses and game show winnings require some personal skill, knowledge, and effort. Tax rebates sometimes differ from the other sources of payment because the tax rebate is a return of the taxpayer's withholdings. That is, outside of refundable credits tied to specific performance, respondents generally cannot materially profit from a tax rebate because it is a return of the taxpayer's own money already paid in but can profit from a game show winning or bonus. Some political rhetoric frames taxes as money belonging fundamentally to taxpayers, not the government, whereas bonuses and game show winnings come with less of an entitlement. Bonuses are likely to be closely tied to an individual's performance, however. Game show winnings might be as well, if the winner attributes success to having a higher skill level than fellow contestants. In addition, collecting a bonus or a tax rebate may be repeatable. One could not count on or control repeating a game show winning.

Additional differences in affective "euphoria" surround these payments. It is unlikely that there will be a TV commercial asking, "you just got a tax rebate, what are you going to do?" "I'm going to Disney World!" However, winning a game show, or perhaps even earning a bonus may be cause for celebration. If the mental frame of the windfall is celebratory, then

perhaps the spending will be directed more toward celebrations and fun than if the windfall was from a tax rebate. Alternatively, if the recipient were looking to brag about or show off their good fortune, they may be more likely to spend it conspicuously on fun than on regular household expenses. They may allocate more toward an infrequent expense such as a vacation, bigger holiday gifts, or something they have always wanted. Differences in the amount spent by classification and by source are to be expected, but no source is absolute and completely separate in characteristics from the other sources, biasing against finding any differences.

Basically, our hypothesis is that the more euphoric and hedonistic the source, like game show winnings, the more one would spend on fun. Alternatively, the more adult the source of the windfall, such as a tax rebate, the more one would allocate to more "responsible" uses, like investing in stocks and bonds, or household expenses and durable goods, such as a car or washing machine. In testing these hypotheses, the amount of the income in dollars and relative to household income, the amount of the payment, the respondent's habit of spending or saving, the order of presentation, the frequency of payments and the demographic characteristics of the respondents will be controlled for.

With that in mind, the null hypotheses are:

Hypothesis 1: There will be no difference in spending on "fun" by source of windfall.
Hypothesis 2: There will be no difference in allocations for regular expenses, credit card payments, durable assets or investing in stocks, bonds and savings account (adult uses of funds) by source of windfall.

Additionally, if either of these hypotheses produce significant findings, the sensitivity of the spending pattern will be analyzed to answer the research question:

RQ_1: Is the amount spent on fun or adult uses bounded at a fixed dollar amount or a relative percentage of the amount received, or is it relatively elastic?

This research question, we believe, has been previously unexplored in research literature and represents a contribution to the knowledge of the field.

4. Method

This study examined respondents' intended uses of hypothetical windfalls. Sheppard et al.'s (1988) meta-analysis of 86 theory-of-reasoned-action studies found a 0.53 correlation between intention and behavior, indicating that intent is a strong predictor of action.

In this study, the intended spending/saving patterns of respondents were gathered through 80 different paper-and-pencil instruments through students' professors at seven universities. Professors familiar with this type of research gathered the responses, sometimes providing a negligible amount of extra credit, and returned the responses to the authors. Each participant was given one of these 80 instruments at random, which contained identical questions except for the source of the income and the amount of the hypothetical cash transferred, and asked how she would use the funds, both if it were received as a lump-sum and if the same amount were received spread out over 12 equal monthly payments (within-subject design) from two

of these five sources: bonus, game show winnings, inheritances, lottery winnings, and tax rebates (between-subjects design). The amount of payments on the instrument was one of these four different amounts: $300, $600, $1,500, and $3,000. Some instruments presented the lump-sum amounts first and some presented the periodic amounts first to control for the order effect.

Consistent with Chambers and Spencer (2008), the instruments asked how much of a lump sum rebate would be used for: (1) investing, (2) paying off credit card debt, (3) paying off notes, (4) regular monthly expenses, (5) buying a durable asset, (6) saving for an infrequent expense, and/or (7) used for fun. The instrument also asked how much of a monthly payment, equal to one-twelth of the lump sum amount, would be used for each of these seven purposes. Similarly, the flip side of each instrument asked these same questions, changing only the source of the payment from one source to another—such as from a tax rebate to a game show payment.

Students were considered provisionally acceptable respondents per Walters-York and Curatola (1998) and Ashton and Kramer (1980). As such, instruments were distributed to university students at these institutions: Coastal Carolina University, Francis Marion University, Longwood University, Metropolitan State University of Denver, Texas A&M University – Corpus Christi, University of Alabama – Birmingham, and University of Houston – Clear Lake. Notably, at least four of these universities have a large nontraditional student population which adds external validity to this study beyond that expected from a traditional student population examined in the academic studies just listed.

All research questions were analyzed with descriptive statistics, and then were converted to a percentage of the total payment received for each of the seven categories: (1) investing, (2) paying off credit card debt, (3) paying off notes, (4) regular monthly expenses, (5) buying a durable asset, (6) saving for an infrequent expense, and (7) used for fun were coded as spending. Because the examples listed in category (6) were "a vacation, bigger holiday gifts, or something you've been wanting," the percentages for items (6) and (7) were added together as were the dependent variable for "Fun spending." The independent variables were Log of income, Materiality of payment, Spending default (that is whether the respondent habitually saves or spends unexpected money received), dummy variables for the total amount of payment, and dummy variables for the source of the windfall: game show winnings, bonus, or tax rebate. Demographic and other control variables were included to control for order effect, risk-taking variables gender, age, business experience level, and education level.

The complete regression model was of the form:

> Percent Spent on Fun = F(Income, Zero income, Amount, Education, Gender, Age, Importance, Seatbelt use, Smoker, Spend1, Experience level, dummy variables for the Source of the payment (tax rebate, bonus, or game show), and a dummy for the Order of presentation (monthly payment first, or lump sum payment first)).

"Income" is the log of the respondent's income plus one. "Amount" is the hypothetical amount of the distribution in dollars. Rather than use a continuous variable for the total payments, dummy variables were created for the four discrete payment amounts. Education was divided into four categories: high school, associate degree, undergraduate degree, and

graduate degree. "Gender" was a categorical male/female variable, where female was coded as "1." "Age" was the participant's age in years. "Materiality" was defined as the total payment divided by the income of the respondent. The "Seatbelt" and "Smoker" dummy variables were included as proxies for respondents' risk preference; seatbelt wearers and smokers were coded as "1." The "Spend1" variable is a measure of respondents' habit of using extra money; the respondents were asked "When you get 'extra money,' do you spend it or save it?" For those that answered "spend," the dummy was set to 1. In testing these hypotheses, the order of presentation and the frequency of payments were also controlled for. Interaction effects were also run as a control measure. Ultimately, the monthly payments were considered immaterial and dropped from the model.

Basic regressions were run matching the two extremes of (un)fun: tax rebate and game show winnings, but eliminating nonsignificant control variables except for Income, Materiality, Spend 1 and Level of payment. That model is:

Hypothesis 1: Percent Spent on Fun = F(Income, Materiality, Spend 1, dummy variables for Amount, and a dummy variable for game show).

Hypothesis 2: Percent Spent on Adult Sources = F(Income, Materiality, Spend 1, dummy variables for Amount, and a dummy variable for game show).

5. Results

There were approximately 1,800 returned instruments in total, of which 601 were usable and pertained to the tax rebate, bonus and game show sources of income. Most of the remaining instruments measured responses for inheritances and lottery winnings, which were not used in this analysis. Some of the instruments were not sufficiently completed, perhaps because some students were trying to get extra credit without doing the work, and because responses were anonymous, turning in a partially completed instrument would produce that effect. If we are correct in reading this situation, however, that would bias this study against findings because of the noise introduced in hastily completed instruments. Twenty-two of these observations had at least one missing value. The results of the regression are shown in Table 1.

As shown in Table 1, the results of this regression were highly significant at $p \leq 0.0005$, although the R^2 is 0.0652 and the adjusted R^2 is 0.0481. Likewise, the source of the payment was highly significant at $p \leq 0.0151$ and the coefficient was a positive 0.06933, indicating that respondents spent more on fun when they received the same amount of payment from a fun source (game show) than from a less fun source (tax rebate), and rejecting the null hypothesis 1. Materiality, which is the relative size of the total payment, was also statistically significant at $p \leq 0.0106$, however, the coefficient of 0.00009549 is economically quite low. The total amount of the payment for each dummy variable was significant at $p \leq 0.05$, with all amount coefficients being negative, indicating that the higher the payment, the less was spent on fun. Spend1, which was the dummy variable equal to one for those that indicated that they would normally spend extra funds, was significant at $p \leq 0.05$ as this would be a

Table 1 Regression of hedonistic "fun" spending between game show and tax rebate

Analysis of variance						
Source	*df*	Sum of squares	Mean square	*F* value	Pr > *F*	R^2
Model	7	2.03636	0.29091	3.82	0.0005	0.0652
Error	383	29.18922	0.07621			
Corrected total	390	31.22558				

Parameter estimates					
Variable	*df*	Parameter estimate	Standard error	*t*-value	Pr > *t*
Intercept	1	0.22191	0.06037	3.68	0.0003
Lnincome	1	0.00778	0.00583	1.33	0.1829
Materiality	1	0.00009549	0.00003720	2.57	0.0106
Spend1	1	0.05509	0.02934	1.88	0.0611
Level600	1	−0.10282	0.03625	−2.84	0.0048
Level1500	1	−0.09549	0.04242	−2.25	0.0249
Level3000	1	−0.13641	0.04201	−3.25	0.0013
Gameshow dummy	1	0.06933	0.02839	2.44	0.0151

one-tailed test for this variable. The results of this regression indicate that the first null hypothesis was rejected.

In the combined, three-source regression shown in Table 2, the model continues to be highly significant at $p \leq 0.01$. The amounts of the payment continue to be highly significant with negative coefficients, and the game show source dummy continues to be highly significant and results in higher spending on fun. When comparing payments from bonus and tax rebates (the omitted variable), the source of the payment was not marginally significant at $p \leq 0.10$, indicating that a bonus was more neutral than either tax rebate or game show sources.

Table 2 Regression of hedonistic "fun" spending among bonus, game show, and tax rebate

Analysis of variance						
Source	*df*	Sum of squares	Mean square	*F* value	Pr > F	R^2
Model	8	1.97917	0.24740	3.21	0.0014	0.0431
Error	570	43.97676	0.07715			
Corrected total	578	45.95592				

Parameter estimates					
Variable	*df*	Parameter estimate	Standard error	*t*-value	Pr > *t*
Intercept	1	0.27961	0.05228	5.35	<.0001
Lnincome	1	0.00184	0.00486	0.38	0.7046
Materiality	1	0.00004837	0.00003032	1.60	0.1111
Spend1	1	0.03874	0.02435	1.59	0.1121
level600	1	−0.08007	0.03120	−2.57	0.0105
level1500	1	−0.11248	0.03409	−3.30	0.0010
level3000	1	−0.11368	0.03441	−3.30	0.0010
Bonus dummy	1	0.03351	0.02911	1.15	0.2502
Gameshow dummy	1	0.06745	0.02840	2.38	0.0179

Table 3 Regression of percent invested in stocks, bonds, and savings comparing bonus, game show and tax rebate

Analysis of variance						
Source	*df*	Sum of squares	Mean square	*F* value	Pr > *F*	R^2
Model	8	4.23598	0.52950	7.51	<.0001	0.0953
Error	570	40.19711	0.07052			
Corrected total	578	44.43309				

Parameter estimates					
Variable	*df*	Parameter estimate	Standard error	*t*-value	Pr > *t*
Intercept	1	0.26953	0.04999	5.39	< .0001
Lnincome	1	−0.00043260	0.00465	−0.09	0.9259
Materiality	1	−0.00002566	0.00002898	−0.89	0.3764
Spend1	1	−0.15297	0.02328	−6.57	< .0001
level600	1	0.04130	0.02983	1.38	0.1667
level1500	1	0.08352	0.03259	2.56	0.0106
level3000	1	0.08799	0.03290	2.67	0.0077
Bonus dummy	1	−0.06468	0.02783	−2.32	0.0205
Gameshow dummy	1	−0.05112	0.02715	−1.88	0.0602

Regressions were also run directly comparing spending on fun from game show winnings with a bonus. No statistically significant differences were found. Similarly, regressions were run directly comparing spending from a tax rebate and a bonus. No statistically significant differences were found and tables for these results are omitted. This may mean that a bonus has characteristics of both a tax rebate and a game show winning. Like a tax rebate, it is derived from work, but like a game show winning, a bonus may have euphoric qualities one would celebrate. Though not statistically significant from zero, the coefficient for Bonus is positive and about half the size of the coefficient for the game show dummy. In the end, while some of a tax rebate would be spent on fun, the regression results indicate that the amount spent on fun from a bonus is not different from the amount spent on fun when the source is either a tax rebate or a game show winning.

Regressions were also run to see if spending on adult uses would differ by the source of the income. Various definitions of "Adult uses" to mean (1) spending on regular monthly expenses, or (2) the sum of regular monthly expenses and paying off credit cards, or (3) the sum of regular monthly expenses, paying off credit cards, and to buy a durable asset (such as a car, boat, washing machine, or furniture) were used. Regardless of the form of the measure used for "Adult spending," none of these regression models produced results significant enough to reject the second null hypothesis and are not presented as a table. However, when regressions were run to see if saving (Investing in stocks, bonds, savings accounts, etc.) increased significantly when the source was a tax rebate instead of a game show, the results were significant (see Table 3). Recipients of tax rebates would allocate more to this type of savings than they would if they received the same windfall amount from either bonuses or game show winnings, confirming that at least to some extent, people tend to use fun sources of income for hedonistic uses and adult sources of income for adult, utilitarian uses, consistent with O'Curry and Strahilevitz (2001). Like O'Curry and Strahi-

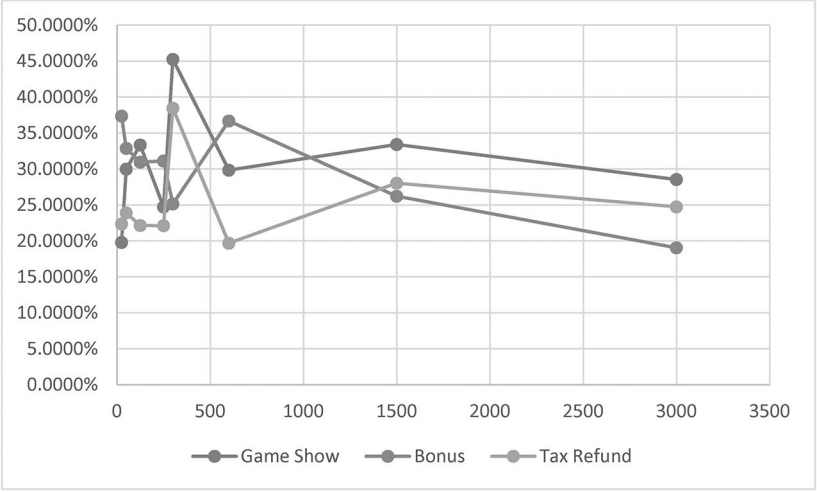

Fig. 1. Percent of windfall spent on fun.

levitz (2001), our amounts varied in value; however, unlike O'Curry and Strahilevitz (2001), we did not ask respondents to assume complete financial independence and allowed respondents to allocate money rather than choose among a selection of prizes that we believe results in a more comprehensive and divisible allocation.

Income in both absolute terms and relatively (as measured by Materiality) were extremely small and insignificant, indicating that the tendency to save and invest transcends income levels, but is strongly dependent on the respondent's savings habits, as indicated by the Spend1 variable.

We then examined the pattern of responses further: were the coefficients for various levels of hedonistic spending linear by source, or did they display a different pattern? When looking at the coefficients for each Level in the Game Show/Tax Rebate model and the significance of the Materiality variable in Model 1, the incremental amount spent appeared to be both significant and nonlinear. To confirm, the average percentage spent on fun was calculated for the monthly amounts of $25, $50, $125, and $250 and the yearly amounts of spending on fun. Next, the average percentage spent on fun for game show, bonus and tax rebate windfalls was graphed. The results, shown in Fig. 1, indicate that for small rebates, the percentage spent on rebates varied, and varied by source. For larger rebates of $1,500 and $3,000, spending on fun leveled out and began to converge at around 30%, regardless of source and then began to slowly fall.

6. Discussion

Overall, this model lends significant support to O'Curry and Strahilevitz (2001) findings of people placing affective tags on money and expands the body of knowledge that one affective tag is fun. Generally, these findings also support Thaler's (1999) mental accounting theory. However, the size of the effects also supports the neo-classical

economic notion that money is more fungible than not, and/or people are more rational than not with their money when it comes to fun sources and fun uses. The effect of affective tags may be bounded.

Regression results presented in Table 2 suggest that survey respondents did not spend the windfall differently if the source was a tax rebate or a work bonus. Additionally, the regressions comparing only windfalls from a bonus and a game show did not show a significant difference between the source. These results seem to indicate that there may be a hierarchy of fun sources. Game shows winnings are likely more fun than work, and work is not much different from taxes, but game shows are clearly more fun than taxes.

Regression results presented in Table 3 suggest that more of tax rebates, which are likely more predictable than bonuses and especially game show winnings, are invested, indicating that clients might be open to making investments during a predictable tax season, providing a greater demand for astute financial advisers like Certified Financial Planners and Certified Public Accountants. Combined with other academic literature on the anticipation of a receipt discussed in the Literature Review section of this article, early, increased communication, especially in the February to April "tax season" might be beneficial to addressing and servicing clients' financial needs. Advisors that are already tax professionals may be at an advantage in serving clients because they know the timing of the receipt, and the amount as well. They also would have the means, with client permission, to split a direct deposit of a tax refund among up to three different accounts with up to three different U.S. financial institutions. Splitting the refund can be accomplished electronically or through the IRS' Form 888, *Allocation of Refund (Including Savings Bond Purchases).*

We found the results of the research question enlightening. We know that respondents have separate mental accounts, or "buckets" (Thaler, 1999). We know those accounts can get full (Chambers, Spencer, and Mollick, 2009). This appears to be what is happening through roughly the $600 payment level. As income rises, so does lifestyle, *ceteris paribus*. However, not all uses of income necessarily rise proportionately. For example, if one's income doubled, that person would not necessarily incur twice as much in medical expenses. A similar increase in income might result in more than doubling a household's federal income tax bill because federal income tax rates are progressive.

Therefore, how do the allocations for fun change with an increase in income? Apparently, at small amounts of affectively tagged windfalls, enough money is spent to fill the current bucket for fun, and then the size of the bucket increases proportionately. The first part of this graph, then, suggests that people can have "enough fun" for their standard of living, confirming Chambers et al. (2009) that buckets get full. The second part of this graph describes the elasticity of fun as windfalls increase, which is an important contribution to literature, which we believe has not yet tested how the components of allocating income, and in particular fun, shift, if at all, with respondents' increase in income. This leads to several questions for further study.

How does the allocation of income, and in particular income from fun sources, shift, if at all, with respondents' increase in income? It appears that while affective tags can produce significant results, hedonistic spending from an affectively tagged source may be, if not

absolutely bounded, relatively bounded. People's rationality, more than not, seems to keep exuberance in check.

Additionally, if people use adult sources like tax rebates for adult uses, then financial professionals can incorporate these findings into their financial advising practices. Because tax filing is an at least annual event, and because most taxpayers receive refunds, tax season may be a robust time for financial professionals to encourage saving from a windfall. To enhance this practice, financial professionals might consider encouraging savings from the current refund and also the coming year's tax refund. From what we know about the power of commitment, those who commit to saving in the future save less than what they commit to saving, but more than those who did not make a commitment at all (Thaler and Sunstein, 2009). Then remind clients of their future commitment throughout the year. Mullainathan and Shafir (2013) showed that by sending a monthly reminder to save by either text or letter, for example, savings increased 6%. This method of future client commitment in the tax setting combined with reminders has not been tested though, so questions still remain, which leads to several questions for further study.

7. Limitations and opportunities for future research

This research showed that there might be a limit to how much of a windfall people are willing to spend on fun, even if the windfall is from a fun source. Future research could explore the elasticity or shape of the spending on fun. Other questions also lend themselves to further research: had windfalls increased further, would the percentage of income allocated to fun stay relatively flat? To what extent is hedonistic spending bounded when the source of the money is affectively tagged as either fun or adult? Had windfalls increased further, would the percentage of income allocated to fun stay relatively constant? Do other allocations of income to, for example, monthly expenses and investments also grow proportionately, or do some level off or even reverse? What are the other affective tags? How does tagging affect income allocations currently, and as the amount of windfall income rises?

One limitation of this article is that it focuses on the changes in behavioral intent when presented with modest windfalls from different sources, and does not examine the latent mental processes (or lack thereof) that are used to reach that intent. We do not disentangle the stimulus, or priming, from the mental accounting that produces the behavioral intent. Priming can be exhibited through what Thaler and Sunstein (2009) would call a "nudge," for example when setting up certain financial defaults to encourage individuals to save for their retirement. Mental accounting on the other hand is an internal construct, but it is connected to nudges that others may use in the environment to improve the choices of people who process information through their mental accounting systems. That is, priming is a cause that when processed with another's mental accounting system may yield a different behavior than that displayed by those who were not primed. In this particular study, we are less concerned with the nuance of disentangling the prime from the latent mental accounting, and more concerned with the type of stimulus and differences in intent.

In this instrument, we asked if respondents smoked and if they wear seatbelts as a proxy for risk aversion. We also asked them for the extent of their business experience. We had seen these questions in previous studies, sometimes in multiple studies. In hindsight, these questions were too general to yield meaningful results. For example, an item on "personal financial expertise" would likely have yielded more information than the more general "business experience."

Similarly, we analyzed the differences in uses among hypothetical receipts from the following sources: game show earnings, bonuses, and tax rebates. Underlying this analysis is that winning money on a game show is more fun than doing one's taxes. That might not be universally true, and that assumption biases against results in this this study. For future research, it might be useful to test the extent that respondents find game show winnings to be more fun than receiving a tax rebate. For example, some people may be so thrilled that they are (hypothetically) getting *any* money that the source is irrelevant, and a separate analysis of those respondents may result in further interesting, significant findings. Another way to examine the effects of fun earnings relative to adult earnings would be to ask respondents what categories of activities they saw as fun rather than adult. These responses could be incorporated into fun or adult scenarios and, among subjects design, could test for different uses of those funds. Such a design may be a better test of these concepts and produce stronger results.

8. Conclusion

Our findings are consistent with the idea that fun sources of income are more likely to be spent on a fun expenditure. Money won on a game show would be spent more on fun than money received from a tax rebate. This provides support for rejecting the first null hypothesis that there will be no difference in spending on fun by source of windfall. However, there may be a hierarchy of fun sources: game shows winnings are likely more fun than work, and work is not much different from taxes, but results show that game shows winnings are clearly more fun than tax rebates.

While some of the regression results were unable to reject null Hypothesis 2 as it pertains to adult spending, regression results for additional investing provide support for rejecting null Hypothesis 2, that there is no difference in allocations for regular expenses, credit card payments, durable assets, or investing in stocks, bonds and savings account (adult uses of funds) by source of windfall. Though various combinations of "adulting" (spending on adult causes) were used, we were unable to show that more adult sources of income are spent on adult *spending* like paying down a credit card or paying regular household expenses. However, there is significant evidence that there is a difference in investing based on the source of the windfall, validating a specific kind of adulting.

Finally, the percentage of the windfall spent on fun levels out. People will apparently spend significantly more on fun when a fun windfall is received, but that spending on fun is not limitless. Additionally, as the amount of the total payment increased, the percentage spent on fun appears to level, indicating that at least within this range of payments, there may be such a thing as "enough (spending on) fun."

Acknowledgment

The authors wish to acknowledge and thank Stetson University for supporting Eugene Bland's development leave.

Appendix Sample Survey Instrument

"What would you do if . . .?" (Fill in the amounts): By participating in a game show, you won a prize that would result in *you* receiving $600.00 for 2012.

If received, how much of these winnings would you plan to:

1. Invest (in stocks, bonds, savings account, and so forth)?	$
2. Use to pay off credit card debt?	$
3. Use to pay off notes (such as mortgage, car note, and so forth)?	$
4. Use up about evenly every month for expenses? _____/month. × 12 months. =	$
5. Use to buy a durable asset (such as car, boat, washing machine, furniture)?	$
6. Use to save for an infrequent expense (such as a vacation, bigger holiday gifts, or something you've been wanting)?	$
7. Spend right away on something fun?	$
Amount must total $600.00————-→	

If instead, by participating in a game show, you won a prize that would result in *you* receiving $50.00/month for the next 12 months.

If received, how much of this monthly increase would you plan to:

8. Invest (in stocks, bonds, savings account, and so forth)?	$
9. Use to pay off credit card debt?	$
10. Use to pay off notes (such as mortgage, car note, and so forth)?	$
11. Use up for regular monthly expenses?	$
12. Use to buy a durable asset (such as car, boat, washing machine, furniture)?	$
13. Use to save for an infrequent yearly expense (such as a vacation, bigger holiday gifts, and/or something you've been wanting)?	$
14. Spend right away on something fun?	$
Amount must total $50.00————-→	

Please list your: Zip Code_____ Years of work experience ____
Highest education level: High School ___ Associate Degree ___ Undergraduate ___ Graduate or above ___
Occupation: _____ Gender: Female ___ Male___ Age ____
Race/ethnicity _____ # of College-level Accounting classes completed College major (if applicable) _____
Industry where you work _____
Approx. yearly ***Household*** *income (from all wage and salary earners and other sources of income) $_____*
Credit Card Debt: $_____ Other Debt: $_____
Do you smoke? Yes ___ No ___ Do you normally wear your seatbelt? Yes —___ No ___
When you normally get "extra money," do you spend it or save it? Spend ___ Save ___
I rate my level of business experience as:
High ___ Fairly High ___ Moderate ___ Fairly Low ___ Low___ None ___

Complete other side, please.
THANK YOU FOR YOUR PARTICIPATION!!!

"What would you do if . . .?" (Fill in the amounts): You got a bonus at work that would result in *you* receiving $600.00 which for 2012 will automatically be mailed to you as a check from your employer.

If enacted, how much of this monthly increase would you plan to:

15. Invest (in stocks, bonds, savings account, and so forth)?	$
16. Use to pay off credit card debt?	$
17. Use to pay off notes (such as mortgage, car note, and so forth)?	$
18. Use up about evenly every month for expenses? _____/month. × 12 months. =	$
19. Use to buy a durable asset (such as car, boat, washing machine, furniture)?	$
20. Use to save for an infrequent expense (such as a vacation, bigger holiday gifts, or something you've been wanting)?	$
21. Spend right away on something fun?	$

Amount must total $600.00————-→

Another work bonus would result in *you* receiving $50.00/month after taxes; that is, your paychecks would go up $50.00/month.

If received, how much of this monthly increase would you plan to:

22. Invest (in stocks, bonds, savings account, and so forth)?	$
23. Use to pay off credit card debt?	$
24. Use to pay off notes (such as mortgage, car note, and so forth)?	$
25. Use up for regular monthly expenses?	$
26. Use to buy a durable asset (such as car, boat, washing machine, furniture)?	$
27. Use to save for an infrequent yearly expense (such as a vacation, bigger holiday gifts, and/or something you've been wanting)?	$
28. Spend right away on something fun?	$

Amount must total $50.00————-→

THANK YOU FOR YOUR PARTICIPATION!!!

References

Abeler, J., & Marklein, F. (2016). Fungibility, labels, and consumption. *Journal of the European Economic Association, 15,* 99–127.

Adamopoulou, E., & Zizza, R. (2017). *Regular Versus Lump-Sum Payments in Union Contracts and Household Consumption.* Working Paper Series 2013, European Central Bank.

Agarwal, S., & Qian, W. (2014). Consumption and debt response to unanticipated income shocks: Evidence from a natural experiment in Singapore. *American Economic Review, 104,* 4205–4230.

Ariely, D. (2008). *Predictably Irrational.* New York, NY: HarperCollins.

Arkes, H. R., Joyner, C. A., Pezzo, M. V., Nash, J. G., Siegel-Jacobs, K., & Stone, E. (1994). The psychology of windfall gains. *Organizational Behavior and Human Decision Processes, 59,* 331–347.

Ashton, R. H., & Kramer, S. S. (1980). Students as surrogates in behavior accounting research. *Journal of Accounting Research, 18,* 1–15.

Baker, M., Nagel, S., & Wurgler, J. (2007). The effect of dividends on consumption. *Brookings Papers on Economic Activity, 1,* 231–276.

Baker, H. K., & Nofsinger, J. R. (2002). Psychological biases of investors. *Financial Services Review, 11,* 97–116.

Benjamin, D. J. (2006). *Psychology, Preferences, and Markets.* ProQuest Dissertations and Theses, pp. 231–231. (available at http://search.proquest.com/docview/305344915?accountid=7084)

Blinder, A. S. (1981). Temporary income taxes and consumer spending. *Journal of Political Economy, 89,* 26–53.

Bodkin, R. (1959). Windfall income and consumption. *The American Economic Review, 49,* 602–614.

Boylan, S. (2010). Prior audits and taxpayer compliance: Experimental evidence on the effect of earned versus endowed income. *Journal of the American Taxation Association, 32,* 73–88.

Bradford, T. W. (2008). Intergenerationally gifted asset dispositions. *Journal of Consumer Research, 36,* 93–111.

Browning, M., & Collado, M. D. (2001). The response of expenditures to anticipated income changes: Panel data estimates. *The American Economic Review, 91,* 681–692.

Camerer, C., Babcock, L., Loewenstein, G., & Thaler, R. (1997). Labor supply of New York City cabdrivers: One day at a time. *Quarterly Journal of Economics, CXIII (May),* 407–442.

Chambers, V., & Spencer, M. (2008). Does changing the timing of a yearly individual tax refund change the amount spent vs. saved? *Journal of Economic Psychology, 29,* 856–862.

Chambers, V., Bland, E., & Spencer, M. (2017). Does the source of a cash flow affect spending versus saving? *Financial Services Review, 26,* 291–313.

Chambers, V., Spencer, M., & Mollick, J. (2009). Goldilocks rebates: Complying with government wishes only when the rebate amount is "just right." *Journal of Economics and Economic Research, 10,* 101–119.

Cheema, A., & Soman, D. (2006). Malleable mental accounting: The effect of flexibility on the justification of attractive spending and consumption decisions. *Journal of Consumer Psychology, 16,* 33–44.

Chen, H., & Volpe, R. P. (2002). Gender differences in personal financial literacy among college students. *Financial Services Review, 11,* 289–307. (available at https://manowar.tamucc.edu/login?url=http://search.proquest.com/docview/212012102?accountid=7084)

Epley, N., Mak, D., & Idson, L. (2006). Bonus or rebate?: The impact of income framing on spending and saving. *Journal of Behavioral Decision Making, 19,* 213–227.

Epley, N., & Gneezy, A. (2007). The framing of financial windfalls and implications for public policy. *The Journal of Socio-Economics, 36,* 36–47.

Fisher, P. J. (2010). Black-White differences in saving behaviors. *Financial Services Review, 19,* 1–16.

Fisher, P. J., Hayhoe, C. R., & Lown, J. M. (2015). Gender differences in saving behaviors among low- to moderate income households. *Financial Services Review, 24,* 1–13.

Frederick, S. (2005). Cognitive reflection and decision making. *The Journal of Economic Perspectives, 19,* 25–42.

Friedman, M. (1957). *A Theory of the Consumption Function.* Princeton, NJ: National Bureau of Economic Research.

Heath, C., & Soll, J. B. (1996). Mental accounting and consumer decisions. *Journal of Consumer Research, 23,* 40–52.

Henderson, P. W., & Peterson, R. A. (1992). Mental accounting and categorization. *Organizational Behavior and Human Decision Processes, 51,* 92–117.

Hershfield, H. E., Sussman, A. B., Rourke, L. O., & Bryan, C. J. (2015). Leveraging psychological insights to encourage the responsible use of consumer debt. *Perspectives on Psychological Science, 10,* 749–752.

Hsieh, C. (2003). Do consumers react to anticipated income changes? Evidence from the Alaska permanent fund. *The American Economic Review, 93,* 397–405.

Johnson, D. S., Parker, J. A., & Souleles, N. S. (2006). Household expenditure and the income tax rebates of 2001. *American Economic Review, 96,* 1589–1610.

Karlsson, N., Gärling, T., & Selart, M. (1999). Explanations of effects of prior income changes on buying decisions. *Journal of Economic Psychology, 20,* 449–463.

Kreinin, M. E. (1961). Windfall income and consumption: additional evidences. *The American Economic Review, 51,* 388–390.

Levav, J., & McGraw, P. (2009). Emotional accounting: How feelings about money influence consumer choice. *Journal of Marketing Research, 46,* 66–80.

Meekin, K. E., Tach, L., & Sykes, J. (2015). *It's Not Like I'm Poor: How Working Families Make Ends Meet in a Post-Welfare World.* Berkeley, CA: University of California Press.

Milkman, K. L., & Beshears, J. (2009). Mental accounting and small windfalls: Evidence from an online grocer. *Journal of Economic Behavior & Organization, 71,* 384–394. (available at https://manowar.tamucc.edu/login?url=http://search.proquest.com/docview/230848483?accountid=7084)

Mullainathan, S., & Shafir, E. (2013). *Scarcity: Why Having Too Little Means So Much.* London: Macmillan.

O'Curry, S. (1999). *Income Source Effects*. Unpublished working paper, DePaul University.

O'Curry, S., & Strahilevitz, M. (2001). Probability and mode of acquisition effects on choices between hedonic and utilitarian options. *Marketing Letters, 12,* 37–49.

Parker, J. (1999). The reaction of household consumption to predictable changes in Social Security taxes. *American Economic Review, 89,* 959–973.

Read, D., Loewenstein, G., & Rabin, M. (1999). Choice bracketing. *Journal of Risk and Uncertainty, 19,* 1–3, 171–197.

Rizzo, J. A., & Zeckhauser, R. J. (2003). Reference incomes, loss aversion, and physician behavior. *The Review of Economics and Statistics, 85,* 909–922.

Rucker, M. H. (1984). Allocation of windfall income: A case study of a retroactive pay increase to university employees. *The Journal of Consumer Affairs, 18,* 101–118.

Sahm, C. R., Shapiro, M. D., & Slemrod, J. (2012). Check in the mail or more in the paycheck: Does the effectiveness of fiscal stimulus depend on how it is delivered? *American Economic Journal-Economic Policy, 4,* 216–250.

Shapiro, M. D., & Slemrod, J. (1995). Consumer response to the timing of income: Evidence from a change in tax withholding. *American Economic Review, 85,* 274–283.

Shaprio, M. D., & Slemrod, J. (2003). Did the 2001 tax rebate stimulate spending? Evidence from taxpayer surveys. In J. M. Poterba (Ed.), *Tax Policy and the Economy* (83–109). Cambridge, MA: National Bureau of Economic Research and MIT Press.

Shefrin, H. M., & Thaler, R. H. (1988). The behavioral life-cycle hypothesis. *Economic Inquiry, 26,* 609–643.

Sheppard, B., Hartwick, J., & Warshaw, P. (1988). The theory of reasoned action: A meta-analysis of past research with recommendations for modifications and future research. *Journal of Consumer Research, 15,* 325–343.

Souleles, N. S. (2002). Consumer response to the Reagan tax cuts. *Journal of Public Economics, 85,* 99–120.

Spencer, M. K., & Chambers, V. (2012). National heuristic shift toward saving any form of tax rebate. *Accounting and the Public Interest, 12,* 106–136.

Thaler, R. H. (1999). Mental accounting matters. *Journal of Behavioral Decision Making, 12,* 183–206.

Thaler, R. H., & Sunstein, C. R. (2009). *Nudge: Improving Decisions About Health, Wealth, and Happiness.* London: Penguin.

Trump, R. K., Finkelstein, S. R., & Connell, P. M. (2015). I will risk a stranger's money, but not my own or my friend's money: Effect of proximity of the money source to the self on financial risk-taking. *Marketing Letters, 26,* 501–512.

Walters-York, L. M., & Curatola, A. P. (1998). Recent evidence on the use of students as surrogates subjects. *Advances in Accounting Behavioral Research, 1,* 123–143.

Wertenbroch, K. (2001). *Self-Rationing: Self-Control in Consumer Choice*. INSEAD Working Paper No. 2001/63/MKT.

Winkelmann, R., Oswald, A. J., & Powdthavee, N. (2011, August). What happens to people after winning the lottery? In *European Economic Association & Econometric Society Parallel Meetings* (August 25–29). Oslo, Norway.

Zagorsky, J. L. (2013). Do people save or spend their inheritances? Understanding what happens to inherited wealth. *Journal of Family Economic Issues, 34,* 64–76.

FINANCIAL SERVICES REVIEW

Financial Services Review 28 (2020) 35-48

A portfolio of leveraged exchange traded funds

William J. Trainor, Jr.[a,*], Indudeep Chhachhi[b], Christopher L. Brown[b]

[a]*East Tennessee State University, Department of Economics and Finance, Box 70686, Johnson City, TN 37614, USA*

[b]*Western Kentucky University, Department of Finance, 1906 College Heights Boulevard, Bowling Green, KY 42101-1061, USA*

Abstract

This study demonstrates how a portfolio of leveraged exchange traded funds (LETFs) targeting a unit exposure to their underlying indexes outperforms a portfolio using traditional ETFs while simultaneously reducing downside risk. By extension, a 3x LETF portfolio designed to mimic 2x LETFs outperforms the underlying 2x LETF portfolio. The results are primarily a function of LETFs borrowing short while the investor lends the additional wealth generated from this leverage in one- to seven-year Treasury bonds or similar type of assets. For every one percent earned above the implied borrowing rate, a portfolio of 2x and 3x LETFs outperforms a traditional portfolio by 0.41% and 0.63%, respectively, corresponding roughly to the additional return on the 50% and 67% of the wealth invested in bonds. More than 90% of LETFs outperformance is explained by the borrowing lending differential.

JEL classification: G11; G17

Keywords: Diversified portfolios; Leveraged exchange traded funds

1. Introduction

Leveraged exchange traded funds (LETFs) were first listed in 2006 by Proshares, although leveraged mutual funds have been around since 1993. While Proshares introduced ±2x products, Direxion upped the leverage ante with ±3x funds in late 2008. Because LETFs are designed to return a daily multiple, the constant daily leverage results in uncertain realized

* Corresponding author. Tel.: +1-423-439-5668; fax: +1-423-439-8381.
E-mail address: Trainor@etsu.edu (W.J. Trainor)

leverage over longer periods of time. In general, realized leverage tends to fall over time because of the volatility of returns (Avenllaneda and Zhang, 2010; Carver, 2009; Cheng and Madhavan, 2009; Trainor and Baryla, 2008).

Historically, investors have used margin to create leverage in their investments. However, with LETFs expansion into everything from equity indexes to oil, gold, currencies, and treasuries, it is now possible to create a diversified portfolio of LETFs that mimic a typical investor's portfolio. This can be done by creating unit exposure to the underlying asset classes freeing up wealth to enhance returns. Specifically, 2x and 3x LETFs only require 50% and 33% of investor's wealth to create the same exposure as investing 100% of an investor's wealth in the underlying ETFs or mutual funds. The advantages of using LETFs, instead of margin, are leverage can be theoretically increased to 3x; no explicit interest costs, no possibility of margin calls, and unlike margin, LETFs can be used in most retirement accounts.

However, LETF's leverage is not free as there are implicit lending costs. In addition, LETFs have higher expense ratios and general leverage decay. To the extent an investor seeks to maintain a specific asset allocation, frequent rebalancing is required along with the associated trading costs and taxes on realized gains. A LETF portfolio structured for unit exposure will outperform its traditional counterpart if and only if the return to the invested excess wealth exceeds the implicit financing costs and higher costs of using LETFs.

In contrast to previous research, this study examines how a portfolio composed almost entirely of LETFs can be created to maintain unit exposure to the underlying indexes without increasing the overall risk. George and Trainor (2017) use similar methodology, but where they use a single LETF to demonstrate how it can be used within a portfolio insurance setting, this study uses a portfolio of LETFs to maintain an asset allocation comparable with traditional buy-and-hold portfolios created from mutual funds or ETFs.

In addition, for risk-seeking investors looking to invest in a diversified portfolio with up to 2x leverage, this study examines if 3x LETFs may be used to the same effect. Investing 67% in a 3x LETF is equivalent to 100% in a 2x LETF. If the return to the remaining 33% of the investor's wealth exceeds the additional borrowing cost, the 3x LETF will outperform.

A critical question for the 2x or 3x strategy suggested above is where to invest the freed-up wealth. This study investigates leveraging bonds instead of equity as this will maintain the risk and structural characteristics of the underlying unleveraged portfolio. Because LETFs typically borrow short to attain their exposure, the portfolios investigated in this study are in effect borrowing short to lend long. This strategy should moderately increase expected return and reduce downside risk as most of the portfolio is invested in less risky bonds.

Results suggest a portfolio composed of 2x or 3x LETFs outperforms a portfolio using standard ETFs based on the same underlying indexes. Even in the low interest environment from 2010 to 17 and using an aggregate bond index for the remaining wealth, a portfolio of 2x or 3x LETFs outperforms a portfolio of standard ETFs by 0.9% and 1.8%, respectively, on an annual basis. For risk seekers, using 3x LETFs to mimic the exposure of 2x LETFs outperforms by 1.5% annually during this period. The critical input is the borrowing-lending differential that explains 90% or more of the LETF portfolio's greater return. The results are essentially confirmed with simulated LETF returns since 1946.

The caveat to this type of strategy is LETF portfolios must be rebalanced more often as their initial positions deviate from "optimal" asset allocations even faster than standard portfolios. A 10% barrier threshold defined as a relative deviation of 10% from the initial allocation is used to determine when portfolios are rebalanced. This active management keeps the risk exposure within reasonable bounds while keeping 2x and 3x rebalancing requirements to approximately quarterly and monthly, respectively. Because of tax effects, the strategies outlined above are more suited to qualified nontaxable accounts.

2. Mathematics of LETFs

LETFs magnify the daily return of an underlying index. Because of this constant daily leverage, the realized leverage over any multiday period can be virtually anything and is a function of the daily leverage ratio, time, return, and variance with the latter generally having the largest effect. Realized leverage can mathematically be expressed by:

$$\frac{\mathrm{LR_T}}{\mathrm{XR_T}} = \frac{(1 + \mathrm{XR_T})^{\mathrm{L}} \exp\left(\frac{(L - L^2)\ \sigma^2 T}{2}\right) - 1}{XR_T} \tag{1}$$

where $\mathrm{LR_T}$ is the return to the leveraged fund, $\mathrm{XR_T}$ is the underlying index return, L is the daily leverage ratio, T is time in days, and σ^2 is the standard daily population variance, (Avellaneda and Zhang, 2010; Cheng and Madhaven, 2009).

On average, the variance dominates and realized leverage over time tends to decline. This effect is greater with higher leverage since a daily leverage ratio of 2x multiplies the term $(1 + \mathrm{XR_T})^{\mathrm{L}}$ by $\exp(-\sigma^2\mathrm{T})$ but a daily leverage ratio of 3x multiplies this term by $\exp(-3\sigma^2\mathrm{T})$.[1] Equation (1) is related to volatility drag, which is the difference between geometric and arithmetic average returns and is a major hindrance to LETF returns over extended periods of time. The relationship between the geometric and arithmetic return of any asset is written as:

$$\mathrm{X_{pT}} = \mathrm{X_{RT}} - 0.5\ \sigma_t^2 \tag{2}$$

where $\mathrm{X_{PT}}$ is the geometric return, $\mathrm{X_{RT}}$ is the arithmetic return, and σ_t^2 is the variance of returns.

As an example, assume a daily return and standard deviation of 0.045% and 1.0%, respectively, which roughly corresponds to the S&P 500 market averages from 1946 to 2017. A 3x LETF multiplies these numbers by 3. Thus, the geometric return over a year for the underlying index assuming 252 trading days is 252*0.045% − 0.5*2.52% = 10.08%, and for a 3x this return is 252*0.135% − 0.5*22.68% = 22.68%.[2] For long-term holdings of LETFs, this volatility drag is a major drawback and clearly shows why leveraged funds usually do not return the daily multiple of the underlying index over time.

Expanding on the work of Scott and Watsun (2013), Ott and Zimmer (2016) show the return of an investment with leverage L is:

$$L_{RT} = X_{RT} + (X_{RT} - R_B)(L - 1) - ½(\sigma^2 L^2) \quad (3)$$

where L_{RT} is the leverage return, X_{RT} is the underlying index return, R_B is the borrowing rate, L is the leverage ratio, and σ^2 is the variance of the underlying index.

To reduce volatility drag relative to the underlying index, this study suggests using only a portion of the investor's portfolio to invest in the LETF. This percentage is set so the effective exposure to the index is the same as if the investor invested only in the underlying index. In effect, the volatility of the position in a LETF is no greater than investing directly in the underlying index. This results in additional wealth available to offset the implied borrowing costs and higher expense ratio of the LETFs. In this way, Equation (3) can be modified to account for the return of a portfolio of ETFs or a portfolio of LETFs.

To simplify, assume a portfolio is composed of just one underlying index. Accounting for the LETF's higher expense ratio along with the additional wealth created from using LETFs, Equation (3) becomes the following:

$$L_{PT} = \frac{1}{L}[X_{RT} + (X_{RT} - R_B)(L - 1) - R_{exp}] - \frac{1}{2}\sigma^2 + \left(1 - \frac{1}{L}\right)R_f \quad (4)$$

where L_{PT} is the return to the portfolio using the LETF, R_{exp} is the LETF expense ratio, and R_f is the return to the risk-free asset.[3] If this asset is not risk-free, there is volatility drag to this asset's return which could easily be adjusted for in Equation (4).

The returns are multiplied by 1/L because only a portion of the portfolio relative to the underlying index is invested in the LETF. When L is 1, signifying no leverage, Equation (4) basically reduces to Equation (2). With leverage, the LETF return is multiplied by the inverse of the leverage. For example, with L equal to 3, the investor attains exactly X_{RT} assuming no borrowing costs or expenses. This result attains since one third of an investor's wealth in a 3x is basically the same as if they had invested 100% in the underlying index. With borrowing costs and higher expenses, they are not equivalent. However, the difference is offset by the return in a risk-free or some other alternative asset represented by R_f.

Volatility drag for wealth in the LETF or in the underlying index is the same as both have equal variance. The volatility drag will only become an issue if the percentage in the underlying LETF varies to the point where the effective exposure of the LETF portfolio diverges from a portfolio using the underlying index. With constant rebalancing, this discrepancy can be eliminated or at least effectively managed. Constant or daily rebalancing is infeasible for most investors because of transaction costs but can be mitigated by periodic rebalancing to keep the 1/L ratio relatively constant over time. Lu, Wang, and Zhang (2012) find an investor can assume a 2x/-2x LETF will maintain its leverage ratio for holding periods up to one month. Thus, daily rebalancing is likely not needed.

To calculate the return from using a LETF relative to investing in the underlying index, Equation (2) is subtracted from Equation (4) to attain the following:

$$L_{PT} - X_{PT} = (1 - \frac{1}{L})R_f - 1/L[(R_B)(L - 1) + R_{exp}] \quad (5)$$

Table 1 Portfolio composition using standard ETFs and LETFs

Asset class	ETF	2x LETF	3x LETF
S&P 500	SPY, 30%	SSO, 15%	UPRO, 10%
Mid cap	IJH, 10%	MVV, 5%	MIDU, 3.33%
Small cap	IWM, 10%	UWM, 5%	TNA, 3.33%
International developed	EFA, 10%	EFO, 5%	EURL, 3.33%
Emerging markets	EEM, 10%	EET, 5%	EDC, 3.33%
Real estate	IYR, 5%	URE, 2.5%	DRN, 1.67%
20-year T-bonds	TLT, 5%	UBT, 2.5%	TMF, 1.67%
7 to 10-year T-bonds	IEF, 5%	UST, 2.5%	TYD, 1.67%
Aggregate bonds	BND, 15%	BND, 57.5%	BND, 71.67%

Note: The base percentage in each asset class is set by the initial percentages in the ETFs. Ticker symbols used to calculate results are shown with each percentage.

Equation (5) can be extended to account for multiple underlying asset classes that may be used by an investor. Similarly, one can compare a 3x LETF with a 2x LETF. However, the basic implication of Equation (5) remains the same. If the excess wealth from using LETFs to create a portfolio (the first term in Equation [5]) is greater than the borrowing costs and higher expense ratio of the LETFs, then the return from using LETFs will be greater than the return from using the underlying indexes. For the comparison of two LETFs, the expense ratios will be approximately the same which reduces Equation (5) to whether the additional wealth gained from the higher leveraged LETF earns a return that exceeds its additional financing costs. This study determines whether this is the case.

3. Data and methodology

Because most LETFs were only recently introduced to the market, empirical research is limited. However, there are now 2x and 3x LETFs that cover the main asset categories found in a typical diversified portfolio including small, mid, and large cap equity funds, international funds, short and long-term bond funds, REITs, and a variety of commodity funds including gold, oil, and currencies for those using less traditional portfolios.

To compare portfolio results using tradeable ETFs, Table 1 shows a portfolio of ETFs and their 2x and 3x counterparts along with the effective percentages in each asset. The percentages are based on an investor who has 50% in domestic equity, 20% in international equites, five percent in a REIT, and 25% in bonds. The domestic equity is split between 30% in the S&P 500 with 10% each in mid and small caps. The international equity is split with 10% in developed and 10% in emerging markets. The 25% in bonds is further delineated by five percent in 20+ year T-bonds, five percent in seven to 10-year T-bonds, and the remaining 15% in an aggregate bond portfolio. A similar mix is used by Considine (2006) comparing portfolios of ETFs to mutual funds. The exact percentages are not critical to the results but are created to represent what a typical investor might have.

The portfolio described above can be also be created using LETFs. For an investor using 2x LETFs, only half as much wealth is needed in each LETF to obtain the same amount of

exposure using the underlying ETFs. For 3x LETFs, only a third of the wealth is needed. Thus, to attain 30% exposure to the S&P 500, an investor needs 15% in a 2x S&P 500 LETF or 10% in a 3x S&P 500 LETF.

All the asset classes, except for aggregate bonds, have a corresponding LETF. The aggregate bond asset class is used as the alternative for the excess wealth available when using LETFs to build a portfolio. This results in the 2x and 3x LETF portfolios investing 57.5% and 71.67% of the portfolio, respectively, in a relatively safe bond portfolio. In absolute terms, 80% of the ETF portfolio is at moderate to high risk based on volatility of the underlying assets whereas only 40% and 27% of the 2x and 3x portfolios, respectively, are exposed to equity markets. For comparison of the 3x LETF portfolio to the 2x LETF portfolio, it is assumed the percentage in the underlying 2x LETFs are the same as the underlying index portfolio, that is, 30% is invested in the 2x S&P, and so forth. For the 3x to attain the same exposure as the 2x, 20% is invested in the 3x S&P and so forth.

With limited historical LETF data, additional theoretical LETF returns are calculated for the period before their inception. Because LETFs attain their exposure using a variety of derivative assets including swaps, there are embedded financing costs increasing with leverage (Charupat and Miu, 2014). Based on the methodology of Scott and Watsun (2013), LETF returns are calculated based on data going back to 1946. This shows how LETF portfolios are likely to perform in a variety of market environments including the very high interest rate period during the early 1980s. The equation to calculate daily returns using the S&P 500 LETF as an example is:

$$R_L = L^*R_{S\&P} - R_{exp} - (L - 1)^*R_B \tag{6}$$

where R_L is the daily return to the LETF with a daily leverage ratio of L, $R_{S\&P}$ is the daily return of the S&P, R_{exp} is the daily expense ratio, and R_B is the borrowing rate using the 90-day T-bill rate as a proxy. Strictly speaking, the one-week/month Libor rate should be used, but Libor data begins in 1986 and to remain consistent with sampled returns before this date, the 90-day T-bill rate is used. The 90-day T-bill has a 98% correlation with Libor and averages 0.2% less than Libor. Thus, the borrowing rate is set at the 90-day T-bill yield +0.2%.

The logic behind Equation (6) is a 2x LETF increases exposure by borrowing $1 for every $1 invested. A 3x LETF borrows $2 for every $1 invested. To determine the validity of Equation (6), theoretical daily, monthly, and annual returns are compared with the actual daily, monthly, and annual returns for the LETFs listed in Table 1. All return data are from The Center for Research in Security Prices (CRSP). Monthly return differences as measured by simulated returns minus the LETF returns average 0.01% assuming an additional 1.2% annual expense ratio. Although LETF's average expense ratio is approximately one percent, there are embedded costs associated with derivatives not accounted for. Using a 1.2% expense ratio reduces the average differences for daily, monthly, and annual returns to near zero for the eight asset classes. Monthly differences ranged from −0.07% for the 2x UBT (20+ year Treasury) to 0.17% for 3x emerging markets EET. Thus, Equation (6) appears to approximate returns accurately enough to simulate LETF returns from index data predating LETF's introduction.

Table 2 Portfolio composition using 1946–2017 historical data

Asset class	Portfolio	2x LETF	3x LETF
S&P 500	50%	25%	16.67%
Mid cap	10%	5%	3.33%
Small cap	10%	5%	3.33%
20-year T-bonds	15%	7.5%	5.00%
7 to 10-year T-bonds	15%	7.5%	5.00%
Bond ladder, 1 to 7 years	0%	50%	66.67%

For historical index data, asset classes are defined as CRSP's S&P 500 index, the 2–4 value weighted deciles proxy for a small-cap fund, and 5–7 value weighted deciles proxy for a midcap fund. The 20-year T-bond and an average of seven to 10-year T-bonds proxy for two additional bond funds. The remainder of a LETF's portfolio is invested equally in one, two, five, and seven-year treasury bonds. No reliable daily international data are available before 1991 so the portfolio comparisons are limited to domestic data. Table 2 shows the portfolio weights for the theoretical historical portfolios. As above, when comparing the 3x to the 2x LETF, it is assumed 100% of the portfolio is in 2x LETFs. This requires two-thirds of the 3x LETF portfolio to be invested in the 3x LETFs.

The final question is rebalancing. If it is assumed the weights in Tables 1 and 2 are optimal, the investor must determine to what extent they can deviate from those percentages. A variety of equity variance thresholds are tested to determine when the portfolios need to be rebalanced. For example, a 10% threshold implies an initial 70% equity exposure is rebalanced when the combined equity position breaches 63% or 77%. The effective exposure for the 2x and 3x portfolios is under the same constraint. For the 2x, this means the amount of wealth in equities can only deviate from 35% by $\pm 3.5\%$ without initiating a rebalance. Daily, monthly, and quarterly rebalancing is also investigated.

4. Results

4.1. Rebalancing

To determine how often rebalancing is required to maintain a consistent risk-profile, monthly, quarterly, and allocation thresholds are tested on simulated historical data from 1946 to 2017. Table 3 gives summary statistics using monthly or a 10% variance threshold for rebalancing relative to initial asset allocations described in Table 2. Results for the 2x and 3x are based on their effective exposure.

With monthly rebalancing, a standard portfolio with 70% exposure to equities deviates from 62.65% to 73.44%. However, the 2x LETF portfolio deviates from 43% to 84%, while a 3x LETF portfolio deviates from 27% to 97%.[4] Quarterly rebalancing saw greater extremes as would be expected suggesting a variance threshold must be set for LETF portfolios to keep the risk profile comparable with using standard ETFs.

Using a 10% variance threshold demonstrates improved results in terms of absolute deviations in exposure relative to being invested in the underlying index. Using a 10%

Table 3 Equity portfolio exposure for ETF and LETFs

	Standard ETFs		2x LETFs		3x LETFs	
	Monthly	10% Rebal	Monthly	10% Rebal	Monthly	10% Rebal
Average	70.08%	71.59%	70.44%	71.05%	70.87%	70.94%
Standard deviation	0.98%	2.80%	3.91%	3.10%	6.86%	3.41%
Min	62.65%	62.59%	43.28%	51.22%	26.72%	36.91%
Max	73.44%	77.28%	83.98%	81.03%	96.92%	85.34%

Note: Summary of portfolio exposure for ETF and LETFs from 1946 to 2017 based on monthly and 10% threshold rebalancing. Initial and rebalanced weights are 70% equities, 30% bonds. Weights for the LETFs are effective exposure.

variance threshold gives results approximately equal to daily rebalancing. For portfolios using either 2x or 3x LETFs, maximum exposure to equities is reduced. A 3x still reaches a maximum equity exposure of 85%, but this is the same exposure reached using daily rebalancing. Thus, a tighter variance threshold is not warranted.

In addition, the standard deviation of the exposure to equities is significantly reduced, especially for 3x LETF portfolios. Over 72 years, using a 10% threshold results in 24 rebalances for the standard portfolio, 243 for the 2x, and 651 for the 3x. This results in rebalancing for a standard portfolio, 2x, and 3x on average every three years, quarterly, and 45 days, respectively.

In terms of transaction costs, assuming six trades at $5 a trade is needed at each rebalance, a standard portfolio valued at $100,000 would have 0.01% additional annual expenses, a 2x would have 0.1%, and a 3x would have 0.27%. For taxable accounts, the additional trading required using LETFs will have a greater tax burden that could nullify any excess returns. These tax effect differentials are discussed in the empirical results.

The reported results in this research are based on using a 10% equity variance threshold for rebalancing. It should be noted 24 rebalances over 72 years using standard ETFs might seem small but consider a 20% increase in equities. With an initial $100 portfolio, the value of equities would increase from $70 to $84, but the percentage in equities only increases to $84/$114 or 73.7%. This still does not breach the 10% barrier even with no increase in the bond position. Thus, a relatively large move is required to breach the barrier. A variety of variance thresholds are tested but return differences are not significantly different across portfolios based on rebalancing rules. A 10% threshold effectively controls asset exposure without requiring excessive trading. Thus, only the 10% variance threshold is reported.

4.2. 2010 to 2017 portfolio returns

Table 4 shows the portfolio results for using the ETFs and LETFs shown in Table 1 from 2010 to 2017. The average annual portfolio return using standard ETFs is 10.26%. Using 2x LETFs with unit exposure to the underlying indexes increases this average return to 11.19%, while using 3x LETFs increases this return to 12.06%. In any individual year, the portfolio returns are relatively similar demonstrating the use of LETFs does not result in any discernable increases in risk.

Table 4 Portfolio returns from 2010 to 2017

	ETFs	50% in 2x	33% in 3x	100% in 2x	67% in 3x	BND return
2010	17.11%	19.98%	20.17%	30.57%	34.52%	4.94%
2011	0.42%	2.78%	5.34%	−3.77%	−1.82%	7.92%
2012	13.62%	15.11%	16.62%	25.29%	28.44%	3.89%
2013	17.19%	14.85%	14.03%	34.88%	33.28%	−2.10%
2014	8.43%	10.13%	12.55%	14.29%	16.46%	5.82%
2015	−1.77%	−1.99%	−2.06%	−6.17%	−6.88%	0.56%
2016	9.83%	10.25%	11.54%	16.21%	18.02%	2.53%
2017	17.21%	18.44%	18.31%	34.90%	36.62%	3.57%
Average	10.26%	11.19%	12.06%	18.27%	19.83%	3.39%
Standard deviation	7.56%	7.61%	7.31%	16.28%	16.69%	3.12%

Note: Annual returns for a portfolio of ETFs, 2x LETFs, and 3x LETFS along with the BND ETF returns to show what return the excess wealth from LETF portfolios attained.

As expected, the LETF portfolio returns relative to using standard ETFs rely heavily on the return to the aggregate bond portfolio (BND). In 2013, the BND ETF has a return of −2.10% which leads to LETF portfolios underperforming. In 2015, the BND return is 0.56% also resulting in slight underperformance by LETF portfolios. However, in 2010, 2011, 2012, 2014, 2016, and 2017 when the aggregate bond fund did relatively well, the LETF portfolios outperform by up to five percent over a traditional portfolio (see the 3x LETF portfolio in 2011). On average, even under a near zero interest rate environment over the last eight years, a portfolio of LETFs outperforms a traditional portfolio.

The 100% in 2x and 67% in 3x columns show the returns for doubling the exposure to the underlying indexes. Returns for most years are approximately doubled, but so are the standard deviations. The compounding issue is easily seen in 2011 as the ETF portfolio is slightly positive, but the fully leveraged 2x portfolio is negative. Finally, the portfolio using 3x LETFs to create exposure equivalent to the 2x LETFs outperforms the 2x LETF portfolio by approximately 1.5% annually.

To further break down the performance of LETFs relative to comparable ETFs, Table 5 shows the returns as if 100% is invested in each asset class relative to 50% in the 2x LETF

Table 5 100% effective exposure for ETF and corresponding LETFs

ETF Ticker	Asset Class	ETF Average return	ETF Standard deviation	2x Relative Performance	3x Relative Performance
SPY	Large cap	14.50%	10.15%	−0.99%	−0.67%
IJH	Mid cap	15.11%	12.88%	−0.68%	−1.08%
IWM	Small cap	14.51%	15.31%	−0.88%	−1.23%
EFA	Developed	7.28%	13.78%	−1.49%	−0.84%
EEM	Emerging	5.84%	19.76%	−1.14%	−1.58%
IYR	Real estate	12.78%	11.81%	−0.70%	−0.18%
TLT	20-year	8.20%	15.48%	−0.47%	−0.95%
IEF	7–10 year	4.29%	6.42%	−0.41%	−0.44%
Average		10.31%	13.20%	−0.85%	−0.87%

Note: Comparison of annual results for 100% in the underlying ETF, 50% in a 2x, and 33.3% in a 3x for each ETF or LETF from 2010 to 2017. Remaining wealth in LETFs assumed to earn zero. No statistical return differences between the ETFs or LETFs.

and 33.3% in the corresponding 3x LETF. The remaining wealth for the LETF portfolios is assumed to earn zero. The 10% variance rebalancing rule is applied to each asset class to keep the exposure close to 100% while still being able to estimate beta decay and the higher costs of LETFs.

On average, 50% of a 2x or 33% of a 3x underperforms its 100% ETF counterpart by 0.85% for 2x LETFs and 0.87% for 3x LETFs. For 2x LETFs, the underperformance ranges from −0.41% for 7 to 10-year Treasuries to −1.49% for developed equity. The 3x LETFs underperformance ranges from −0.018% for Real Estate to −1.58% for Emerging Markets. The discrepancy in ranges is due in part to differences in the volatility/return differences across assets and tracking error. It is not surprising the worst underperformance for the LETFs occurs in the volatile international markets that have substandard returns given their volatility. The standard deviations for the ETFs and LETFs are virtually identical and thus, only the ETF standard deviations are shown. In summary, these results show the inherent costs of the higher expense ratio and financing costs of LETFs. From a portfolio standpoint, the remaining funds available from using LETFs need to overcome this performance lag.

To determine how well each LETF tracks their ETF counterpart, LETF returns are also regressed on their corresponding ETF returns from 2010 to 2017. Beta coefficients range from 0.97 to 1.04 with 12 of 16 at 0.99 or 1.0. The 2x emerging market LETF is the only exception with a beta of 0.90 and an r^2 of 90%; r^2 for all other LETF regressions are 98% or better. Thus, of the 16 LETFs, only one did not attain almost exact unit exposure. These results contrast to Tang and Xu (2013) who show even realized daily leverage was less than advertised from 2006 to 2010.

However, that period was more volatile and the LIBOR rate was higher. In addition, small tracking errors are not as magnified in this study as only the inverse of the leverage is held in each fund. The market was also generally upward trending during this study's time period resulting in a positive compounding effect, and finally, the funds themselves are possibly doing a better job of maintaining their daily leverage ratios. Thus, the results suggest using LETFs along with the 10% variance threshold achieves the goal of 100% exposure to the underlying indexes, albeit with an annual financing and expense ratio drag of approximately −0.85%.

4.3. 1946 to 2017 simulated historical returns

To get a better idea of portfolio returns going forward, simulated returns from 1946 to 2017 are created using Equation (6). Table 6 shows the portfolio results using the weighting shown in Table 2. Subperiod returns are also shown roughly corresponding to different interest rate environments. For the entire 1946–2017 period, LETF portfolios on average outperform a traditional portfolio from 0.63% to 1.41% a year.

However, if this strategy is implemented in nonqualified accounts, the differences do not overcome trading costs and taxes. Assuming a long-term capital gains tax of 15% and a 24% marginal income tax bracket, the approximate returns after taxes and trading costs for the 1946 to 2017 time period for the ETF, 2x, and 3x portfolios are 10.54%, 9.75%, and 10.45%, respectively. This assumes all gains to the ETF are taxed at 15% every three years and all

Table 6 Portfolio average annual returns using 1946 to 2017 historical data

Time period	Portfolio	50% in 2x	33% in 3x	100% in 2x	67% in 3x	Treasuries
1946–2017	11.12%	11.75%	12.53%	18.96%	20.58%	5.46%
1946–1959	11.45%	10.78%	11.22%	21.91%	22.23%	1.63%
1960–1978	8.03%	8.45%	9.35%	13.26%	14.82%	5.39%
1979–1991	16.56%	18.81%	20.05%	25.40%	28.75%	11.42%
1992–2009	10.13%	11.00%	11.65%	15.40%	17.22%	5.78%
2010–2017	11.29%	11.47%	12.16%	23.80%	24.72%	1.96%
1946–2017 risk statistics						
Standard deviation	12.67%	13.23%	13.51%	26.52%	26.93%	5.09%
Min	−23.38%	−19.45%	−18.64%	−46.33%	−44.81%	−1.66%
Max	38.84%	39.53%	44.18%	89.92%	90.62%	25.77%
VaR	−5.31%	−4.24%	−3.19%	−14.57%	−12.52%	0.51%
Sharpe	0.55	0.57	0.62	0.56	0.61	N/A
Sortino	2.56	2.93	3.34	1.97	2.26	N/A

Note: Table shows average annual returns from 1946 to 2017 along with several sub-periods corresponding to different interest rate environments. Because of limited data and large standard deviations, there are no statistical differences between the mean returns between the unit exposure portfolios or between the 2x exposure.

gains to the LETFs are taxed at 24% each year. Thus, after taxes and trading costs, using LETFs do not outperform a standard ETF portfolio.

For qualified nontaxable accounts, LETF's outperform, especially when the interest rate environment is relatively high such as 1979 to 1991. Alternatively, during periods of extremely low rates such as the 1946–1959, a portfolio of LETFs tends to underperform. However, low rates by themselves do not relegate LETF portfolios to underperformance. Both the empirical data in Table 4 and simulated data in Table 6 show LETF portfolios outperforming during 2010–2017. The critical factor is the return to the freed-up wealth from using LETFs relative to their implicit financing costs and higher expense ratios.

The bottom of Table 6 shows the standard deviation, minimum, maximum, 10% Value at Risk (VaR), along with the Sharpe and Sortino ratio, the latter of which measures downside risk (Sortino and Price, 1994). Even though there is a marginal increase in the standard deviation when using a portfolio of LETFs, the minimum and 10% VaR for the LETF portfolios is better than the traditional portfolio while at the same time having higher maximums. Thus, the standard deviation is misleading when measuring risk for the LETF portfolios as downside risk for LETFs is mitigated because of the higher percentage in a relatively safe bond ladder. The large percentage in the bond ladder is also valuable during market panics as investors flee to safer assets. Results from the simulated returns reinforce the empirical data from 2010 to 2017.[5]

For investors seeking 2x leverage, the 100% in 2x and 67% in 3x columns show the results. Average returns are less than double while the standard deviations and minimums are doubled with a tripling of the VaR. Using the 3x LETF to attain 2x exposure shows improved results. Overall and in every subperiod, 67% in a 3x outperforms 100% in a 2x with a 1.6% annual average increase. The minimum and VaR using 3x LETFs are also better than 2x LETFs along with a higher Sharpe and Sortino ratio.

Table 7 Return differences regressed on T-bond ladder or AGG ETF minus borrowing rate

	r^2	Intercept (*t*-stat)	Coefficient (*t*-stat)
2010 to 2017			
2x LETF–ETF portfolio	92.53%	−0.001 (−1.87)	0.41 (13.7)a
3x LETF–ETF portfolio	95.36%	0.000 (−1.17)	0.63 (17.36)[a]
3x LETF–2x LETF	64.70%	0.000 (0.35)	0.34 (4.65)[a]
1946 to 2017			
2x LETF–ETF portfolio	93.13%	−0.0071 (−6.14)[a]	0.56 (21.39)[a]
3x LETF–ETF portfolio	95.15%	−0.005 (−3.54)[a]	0.76 (25.88)[a]
3x LETF–2x LETF	79.75%	0.003 (1.15)	0.54 (11.06)a

Note: [a]Statistical difference at the one percent level.

The disadvantage to this strategy is the rebalancing frequency required for a 3x. Using a 10% variance threshold to maintain the initial 2x exposure to the underlying indexes requires 79 and 381 rebalances for the 2x and 3x, respectively, over the 72 years. Thus, although using the 3x to mimic the exposure of a 2x increases the returns, trading costs and the more onerous tax effects in nonqualified accounts eliminate the after-tax return differential.

For a LETF portfolio to outperform, the returns to the T-bond ladder need to exceed the financing and expense ratio costs of LETFs. However, this only occurs 45% of the time from 1946 to 2017. Despite this, 2x and 3x LETF portfolios outperform 54% and 61% of the time, respectively, during this time period. This is possible because a highly positive trending market allows LETFs to return more over time than their daily leverage multiple implies. This can make up for small interest rate spreads. To more accurately estimate the effect of the lending minus borrowing differential on returns, the 50% in 2x and 33% in 3x LETF returns minus the standard portfolio returns are regressed on the bond minus borrowing rate differential for both the 2010 to 2017 empirical data and the 1946 to 2017 simulated data. The regression results are shown in Table 7.

Not surprisingly, the bond return minus the borrowing cost explains more than 90% of the return differences of the 2x and 3x LETF portfolios relative to a standard portfolio for both the empirical and simulated historical data. Based on the empirical data, a one percent return over the borrowing rate leads to a 0.41% increase for a 2x and 0.63% increase for a 3x LETF over a standard portfolio. These regression results held when comparing the LETFs to their corresponding ETFs at an individual level and roughly correspond to the 50% and 67% freed up wealth from using 2x or 3x LETF portfolios, respectively.

Results are similar for the 3x LETF relative to the 2x LETF as a one percent return over the borrowing rate corresponds to a 0.34 percentage point increase using the empirical data and 0.54 percentage point gain for the simulated data. Both results are associated with moderately lower r^2 as there is much more volatility with returns leveraged to 2x. Thus, both the empirical and simulated data confirms the importance of the lending-borrowing spread.

5. Conclusion

Despite the early negative press about the dangers of investing in LETFs, they have become popular investment vehicles with 265+ funds growing to more than $68 billion in

assets over the last 12 years (www.etf.com). While LETFs have higher expense ratios and suffer from leverage decay, LETFs can be used effectively to improve average returns and reduce downside risk if properly managed. One key is to periodically rebalance LETFs to minimize both the impact of decay and maintain a set risk exposure.

This study compares a diversified portfolio of typical ETFs to portfolios comprised of 50% in 2x LETFs or 33% in 3x LETFs with the remainder invested in a relatively safe T-bond ladder or a bond fund. In addition, a 3x LETF portfolio is compared with a 2x LETF portfolio where the underlying exposure of the 3x is set equal to the 2x. Using a 10% variance threshold for rebalancing, this study finds a portfolio using LETFs since 2010 or one based on simulated data from 1946 to 2017 outperforms a standard portfolio. By extension, the 3x portfolio created to mimic the 2x portfolio outperforms the 2x portfolio.

Specifically, combining tradeable LETFs since 2010 with an aggregate bond ETF, a 2x LETF portfolio averages 0.9% more per year while a 3x averages 1.8% more relative to a standard ETF portfolio. The 3x portfolio created to mimic the 2x portfolio outperforms by 1.5% annually. Using simulated data from 1946 to 2017, a 2x LETF or 3x LETF portfolio combined with a one, three, five, and seven-year treasury ladder outperforms a standard ETF portfolio by 0.6% and 1.4%, respectively, on an annual basis with no increase in risk and better downside risk metrics as measured by minimums, VaR, and Sortino ratios. The results extend for the 3x over the 2x with an average annual outperformance of 1.6%.

The critical component for LETF portfolio outperformance is how the return to the remainder of the portfolio not invested in LETFs compares to the implicit borrowing costs and higher expense ratios of LETFs. Regression results show the difference in the bond return minus borrowing rate explains more than 90% of the difference in returns between a portfolio created with LETFs versus one created with standard mutual funds. For every one percent return earned over the borrowing rate, the 2x LETF portfolio outperforms by 0.41% and a 3x LETF portfolio outperforms by 0.63%. This corresponds closely to the freed-up wealth invested in bonds.

From a practitioner's point of view, LETFs require more active management and in any given day, these instruments are more volatile. If the rebalance rule is breached on the downside, an investor will need the "stomach" to buy more of a fund that theoretically could lose 60% or more of its value in a day (recall October 19, 1987 when the S&P 500 fell 20%). Thus, a certain degree of behavioral fortitude may be needed before creating a portfolio of 2x or 3x LETFs. In addition, the tax liability from the greater rebalancing neutralizes the excess returns from using LETFs found in this study. Thus, the implementation of the strategy should probably be limited to qualified accounts. Although tax issues are detrimental, the proposed LETF strategy does have the ability to be used in IRA type accounts—accounts that typically do not allow margin trading.

In summary, this study shows LETFs can successfully be held long-term as major components of a portfolio while improving returns and reducing downside risk. Whether piecemeal such as only using an S&P LETF to partially or fully take the place of an S&P ETF, or to completely replicate an investor's entire portfolio of ETFs, LETFs, when properly managed, can enhance returns and reduce risk.

Notes

1 As an example, a 10% return to the index over a 252-day trading period with a one percent daily standard deviation will result in a theoretical realized leverage ratio for a 2x and 3x LETF of 1.80 and 2.34, respectively.

2 The variance over time is $\sigma_t^2 = t\sigma^2$ so for the index, $\sigma_t^2 = 252*0.01^2 = 2.52\%$. For the 3x, the daily returns and standard deviations are 3*0.045 (.135%) and 3*1 (3%), respectively. Thus, for the 3x, $\sigma_t^2 = 252*0.03^2 = 22.68\%$.

3 Note σ^2 L^2 after dividing σ by L becomes $(\sigma/L)^2$ L^2 that simplifies to σ^2.

4 The max and min exposure levels are for the entire period and do not necessarily sum up to 100%. In addition, for the LETF portfolios, the exposure does not include the allocation to the Treasury ladder meaning the combined equity and bond allocation will deviate from 100% between rebalances.

5 Block bootstrapping of the 1946–2017 data was performed to create 20,000 unique one-year returns. The results are qualitatively like the original historical results with the only differences being the mean returns and VaRs across the portfolios are statistically significantly different.

References

Avenllaneda, M., & Zhang, S. (2010). Path-dependence of leverage ETF returns. *Society for Industrial and Applied Mathematics, 1,* 586–603.

Carver, A. (2009). Do leveraged and inverse ETFs converge to zero? *The Journal of Derivatives, 1,* 144–149.

Charupat, N., & Miu, P. (2014). A new method to measure the performance of leveraged exchange-traded funds. *Financial Review, 49,* 735–763.

Cheng, M., & Madhavan, A. (2009). The dynamics of leveraged and inverse exchange traded funds. *Journal of Investment Management, 7,* 43–62.

Considine, G. (2006). *Comparing Portfolios of ETFs and Mutual Funds, Quantext.* (available at https://docplayer.net/27758586-Geoff-considine-ph-d.html)

George, J., & Trainor, W. (2018). Portfolio insurance using leveraged ETFs. *Financial Services Review, 26,* 1–17.

Lu, L., Wang, J., & Zhang, G. (2012). Long term performance of leveraged ETFs. *Financial Services Review, 21,* 63–80.

Ott, R., & Zimmer, T. (2016). Determining the return-maximizing portfolio leverage and its limitations. *Financial Services Review, 25,* 415–425.

Scott, J., & Watsun, J. (2013). The floor-leverage rule for retirement. *Financial Analysts Journal, 69,* 45–60.

Sortino, F., & Price, L. (1994). Performance measurement in a downside risk framework. *The Journal of Investing, 3,* 59–54.

Tang, H., & Xu, X. (2013). Solving the return deviation conundrum of leveraged exchange-traded funds. *Journal of financial and Quantitative Analysis, 8,* 309–342.

Trainor, W., & Baryla, E. (2008). Leveraged ETFs: A risky double that doesn't multiply by two. *Journal of Financial Planning, 21,* 48–55.

FINANCIAL SERVICES REVIEW

Financial Services Review 28 (2020) 49-65

Are multiple share class funds poorly governed?

Jonathan Handy[a,*], Thomas Smythe[b]

[a]*Department of Business and Accounting, Furman University, 3300 Poinsett Highway, Greenville, SC 29613, USA*

[b]*Lutgert College of Business, Florida Gulf Coast University, 10501 FGCU Boulevard South, Fort Myers, FL 33965-6565, USA*

Abstract

Utilizing independent Morningstar Stewardship Grades, this article finds that multiple share class mutual funds (MS funds) have lower quality governance. Ordered probit regressions indicate MS funds are more likely to have lower board quality ratings and managerial incentive ratings, additional evidence the MS structure has not provided the benefits initially put forth by supporters. The results continue to demonstrate that less sophisticated investors seeking financial advice (those typically utilizing MS funds) may potentially be directed to funds that underperform and have higher costs.

JEL classification: G180; G230; G340

Keywords: Mutual fund; SEC; Corporate governance; Shareholders; Investor protection

1. Introduction

With the 1995 establishment of the Securities and Exchange Commission Rule 18f-3 (Securities and Exchange Commission, 1995), came the widespread use of multiple share class mutual funds (MS funds). MS funds are designed to have a single investment portfolio with a variety of combinations of commission structures and ongoing expense structures. Individuals choose a particular combination from those available, the class, and invest in the fund. O'Neal (1999) points out that such funds are complex and that choosing the most advantageous combination of fund and class may be difficult for investors. In most situations

* Corresponding author. Tel.: +1-864-294-3139.
E-mail address: jonathan.handy@furman.edu (J. Handy)

1057-0810/20/$ – see front matter

where MS funds are utilized, the investment is made with the aid of a financial advisor/broker. However, research shows there is confusion among advisors as to the most appropriate MS fund class based on expected investor holding period (Jones, Lesseig, and Smythe, 2005a). Moreover, recent research finds that MS funds have higher expense ratios than single-class (Non-MS) funds when taking fund governance quality into consideration (Handy, Nichols, and Smythe, 2018). Because fund expenses are directly correlated with investor returns, one might expect advisors/brokers to place great emphasis on them when recommending funds to clients, but research by Jones, Lesseig, and Smythe (2005b) suggests that they rarely do. Because previous research ties fund expenses to governance quality, and prior results indicate that advisors/brokers may not use fund expenses as a major selection criteria, this article investigates the relative governance quality of MS and Non-MS funds so that investors and advisors/brokers are better informed regarding their investment selections.

While motivated by O'Neal (1999), which explores differences across *fund classes*, this article explores differences across *fund structure*. Specifically, this article explores a more subtle issue in the mutual fund market: whether MS funds have differing levels of governance quality compared with Non-MS funds. To do so, we use independent governance ratings data from Morningstar.

Beginning in 2004, Morningstar began publishing Stewardship Grades for mutual funds based on five criteria: board quality, managerial incentives, fees, corporate culture (of the fund sponsor), and regulatory ratings. Each grade provides information on fund management in the context of governance and administrative functionality, and as stated by Morningstar, "helps investors to assess funds based on the degree to which the funds' parent—the management company offering the fund—has its interests aligned with those of fund shareholders." Over the last 10 years, a growing stream of research examines whether the stewardship grade, or its components, have any relationship to fund performance and/or fund costs. To date, this is the first article using stewardship grades to analyze whether there are differences in fund governance across MS and Non-MS funds.

Our analysis utilizes components of the Morningstar Stewardship Grade (MSG) to gauge the degree of governance alignment between retail investors and fund management companies. More specifically, we analyze whether funds selected by investors choosing to use financial advisors have systematically different governance ratings from funds offered to retail investors not using financial advisors.[1] To be clear, we are not suggesting financial advisors consciously recommend funds with differential governance; however, the primary pool of funds from which financial advisors make recommendations does consist of MS funds. In fact, for advisors who work for large securities firms, the pool of funds from which they choose is preselected at the firm level, and so advisors may be forced to suggest such funds. Ultimately, it follows that if MS funds are generally more likely to have lower governance ratings, then retail investors utilizing advisors are more likely to be steered towards funds with lower governance quality and need to know this.

Our findings identify significant differences in Morningstar governance ratings across fund structures. MS funds are more likely to have lower board quality ratings and lower managerial incentive ratings than Non-MS funds (traditional no-load funds). Examining predicted probabilities, the results indicate that MS funds have a significantly lower probability of having high governance ratings when compared with Non-MS funds. Additionally,

a fund sponsor's corporate culture, as rated by Morningstar, has predictable independent influences on fund board quality and managerial incentives. Finally, funds with better board quality ratings have higher managerial incentive ratings. Our findings provide additional evidence that the MS structure is associated with largely detrimental effects on a less informed group of investors.

The balance of the article proceeds as follows. In Section 2, we review mutual fund literature to place this work in context. In Section 3, we develop our hypotheses. In Section 4, we introduce the data, build the empirical model, and discuss the variables of interest. In Section 5, we present primary empirical results, and in Section 6, we present our concluding remarks.

2. Previous literature

Over the years, the use of mutual funds for investing has increased tremendously. Fund assets increased from $11.1 trillion at year-end 2009 to $17.7 trillion at year-end 2018 (2010 and 2019 Investment Company Fact Books, respectively). Accordingly, mutual fund research remains a growing specialty. For brevity, our literature review focuses on the three areas of fund research most related to this article: research focusing on the nature of multiple-share class funds, research focusing on the intersection of retail investors and financial advisors, and research focusing on Morningstar Stewardship Grades (MSG).

2.1. Multi-share class mutual funds

Despite the fact that the SEC has allowed mutual fund sponsors to offer MS funds since 1995, little analytical research into MS funds and the MS structure exists. This article, to the best of its authors' knowledge, is the first to analyze whether there are differences in independently rated governance quality between MS and Non-MS funds. The main theoretical multi-share class research comes from Livingston and O'Neal (1998), O'Neal (1999), and Nanda, Wang, and Zheng (2009). Lesseig, Long, and Smythe (2002) and Handy, Nichols, and Smythe (2018) provide empirical results.

Livingston and O'Neal (1998) and O'Neal (1999) focus on MS fund costs and the incentives provided to investors and brokers.[2] Livingston and O'Neal (1998) concludes that, because little evidence supports mutual fund performance persistence, investors should select funds based on costs because higher costs uniformly lead to lower returns. They identify the most common MS fund cost distribution types, derive a series of mathematical equations expressing the costs as a present value, and provide a comprehensive set of optimal investment strategies for investors given specific investment time horizons. However, when considering the results of Barber, Odean, and Zheng (2005), which demonstrates investor confusion with fund costs, the challenge facing MS fund investors becomes clear: investors must not only choose which fund to invest their assets but also which combination of commissions and ongoing expenses best suits their needs (i.e., they must choose the right class).

O'Neal (1999) derives commission-based incentives for fund brokers/advisors and finds conflicts of interest between advisors and investors. O'Neal (1999) indicates that this is particularly dangerous to MS fund investors, given that those most likely to seek out advisors are those who are relatively uninformed. Nanda et al. (2009) largely focus their analysis around a fund's decision to switch from single-class to the multiple-class structure and find that switching to the MS structure negatively impacts performance. While similar to O'Neal (1999), this article analyzes the relationship between fund *sponsors* and investors by examining Morningstar's independent board quality and managerial incentive ratings to determine if there are differences between MS and Non-MS funds.

Finally, Lesseig, Long, and Smythe (2002) and Handy, Nichols, and Smythe (2018) also focus on MS structure by examining differences in expense ratios between MS and Non-MS funds. Lesseig et al. (2002) analyzes the claim made by fund sponsors when the MS structure was introduced that it allows funds to decrease fund expenses. Their results suggest the opposite—overall net expense ratios for MS funds are higher than for Non-MS funds. Handy et al. (2018) examines a longer and more recent sample and find results consistent with Lesseig et al. (2002).

2.2. Retail investors and the financial advisor

While literature focusing on the MS structure is scarce, literature focusing on the intersection between financial advisors, mutual funds, and the retail investor is abundant. Nofsinger and Varma (2007) analyzes survey results and find that financial advisors are on average more analytical than the general population, that they are more financially patient, and that they perform better in intertemporal choice problems. Bergstresser, Chalmers, and Tufano (2009) analyzes broker-sold (MS) and direct-sold (Non-MS) funds from 1996 to 2004 and do not find any evidence brokers offer substantial benefits to clients.[3] Additionally, Bergstresser et al. (2009) finds that broker-sold funds are no more skilled at aggregate-level asset allocation than funds sold through the direct channel.

While early literature focuses on how individual investors make fund investment decisions (e.g., Alexander, Jones, and Nigro, 1998; Capon, Fitsimmons, and Prince, 1996), Jones et al. (2005a) surveys over 500 financial advisors on what criteria and information sources they use in the fund recommendation process. They find that the two most important information types used are comprehensive data sources and independent rankings from firms such as Morningstar and Lipper (now part of Thomson-Reuters). However, they also find that advisors rank fund costs very low as recommendation criteria. Jones et al. (2005b) also examines survey data from financial advisors regarding their compensation and investment recommendations as it relates specifically to MS funds. They find advisors are more likely to recommend a specific MS class based on the commission received rather than the appropriateness of the class for the client. Additionally, when the funds are firm proprietary funds, advisors are more likely to recommend the class most profitable for the firm, usually to the detriment of investors. The results from Jones et al. (2005b) are consistent with the cautions presented by O'Neal (1999).

2.3. Morningstar Stewardship Grades

MSGs were first introduced in 2004 to "help investors further research, identify, and compare fund managers and fund companies that do a good job—or poor job—of aligning their interests with those of fund shareholders" (Fact Sheet, 2006). In short, the ratings are designed to help investors and advisors evaluate a fund's effectiveness at mitigating the principal-agent problem between investors and fund management. This article brings attention to the grades as an empirical tool, but more importantly, it examines whether there are differences in the governance ratings across fund structure, MS versus Non-MS funds.

Recent work analyzing MSG ratings includes Moore and Porter (2017). They analyze a 2007 cross-section of funds and report that increased mutual fund governance quality, as measured by Morningstar ratings, lead to lower fund expenses. Cao, Ghosh, Goh, and Ng (2014) establishes that MSGs have Granger Causality on long-term risk adjusted returns and can offer an explanation for fund performance, even when Morningstar Star Ratings are considered. Chou, Ng, and Wang (2011) finds that firms with better governance practices, as measured by Morningstar, tend to vote more responsibly on corporate governance proposals of portfolio firms and generally provide better return performance.

This article, as it relates to MSG ratings, is most similar to work by Handy et al. (2018), which examines whether Morningstar's board quality and managerial incentive scores are correlated with fund net expense ratios. Handy et al. (2018) argues that investors should seek to minimize fund expenses and analyze MSG ratings and their relationship to fund expenses as a potential tool for investors to gauge a fund's attractiveness. Of particular interest to the current article is that Handy et al. (2018) find that the relationships between MSG ratings and fund expenses differ between MS funds and Non-MS funds. This article should be considered a more general extension of their work. Rather than focus on MSG ratings in the context of fund expenses, this article looks at the more general question of whether governance ratings differ across distribution channels.

Given the empirical results cited above, the importance of this article should be clear. If lower governance ratings are associated with higher expenses and thereby lower returns, then investors investing in funds with such ratings are being negatively impacted and should be made aware. Consequently, exploring whether or not MS funds have better or worse governance ratings than non-MS funds is a valid pursuit.

3. Hypothesis development

MS funds are targeted primarily to more vulnerable investors, suggesting the need for strong fund-level governance.[4] Thus, our focus is on whether MS funds and Non-MS funds have differences in governance quality as reflected by MSG ratings. When MS funds were introduced, the SEC was concerned about inequitable treatment of shareholders *across* fund classes. However, Handy et al. (2018) shows that MS *funds* have higher net expenses than Non-MS funds and that governance measures have differential effects on expenses across fund structure. As such, our analysis is at the *fund level* and focuses on differences across *fund structure*.

The *Board Qual Rate* and *Manager Incent Rate* variables are evaluated independently. The variables are measured on a scale of 1 (lowest) to 5 (highest). As promoted by Morningstar, each should provide information to investors/financial advisors about the relative quality of fund governance along these dimensions, each of which is important in mitigating the principal-agent problem between fund sponsors and investors.

When estimating the empirical model for each dependent variable, we include the dummy variable, *MS*, equal to 1 if the observation is an MS fund and 0 otherwise. Given that the primary investors in MS funds are considered less knowledgeable, the funds may take additional steps to promote good governance practices. If so, we expect *MS* to have a direct relationship with the governance metrics *Board Qual Rate* and *Manager Incent Rate*. As such,

Hypothesis 1(a): MS funds have higher board quality ratings (*Board Qual Rate*) than Non-MS funds.

Hypothesis 1(b): MS funds have higher managerial incentive ratings (*Manager Incent Rate*) than Non-MS funds.

While examining differences in governance ratings across fund structure is our primary focus, we also are interested in whether the corporate culture of the fund the fund sponsor influences fund governance. As such, we include Morningstar's *Corp Culture Rate* in the empirical models (a variable ranging from 1 [lowest] to 5 [highest]). *Corp Culture Rate* is meant to "assess how seriously a firm takes its fiduciary duty to its fund shareholders." It is an indirect measure of how fund sponsors may influence the governance process *within* funds they operate. We expect more highly rated fund sponsors to have boards that are of higher quality and stronger managerial incentives. As such,

Hypothesis 2(a): Funds whose sponsor has a higher corporate culture rating (*Corp Culture Rate*) have a higher board quality rating (*Board Qual Rate*).

Hypothesis 2(b): Funds whose sponsor has a higher corporate culture rating (*Corp Culture Rate*) have a higher managerial incentive rating (*Manager Incent Rate*).

Finally, once a board is in place, it has the authority to influence contracts between fund managers and the fund as it pertains to managerial incentives. While the board has sole authority to negotiate fund expenses, it is also likely that the board will have an influence on how much fund managers must own to align the interests between the two groups. Therefore, we expect funds with more highly rated boards to have higher managerial incentive ratings.

Hypothesis 3: Funds with higher board quality ratings (*Board Qual Rate*) have higher managerial incentive ratings (*Manager Incent Rate*).

4. Data and empirical model

4.1. Data

The data for the analysis comes from Morningstar and includes year-end data from 2005 to 2009.[5] Our sample only includes funds in the investment objectives Growth and Income,

Growth, Aggressive Growth, and Small Cap for two reasons. First, early mutual fund literature commonly examined these investment categories. Second, and more importantly for this article, the identification of classes in the same fund portfolio had to be identified and coded by hand across all years in the sample, a time consuming process.

Morningstar observations are often referred to as "a fund," but they are not. Morningstar captures data at the *class* level, reflecting differences across share classes that a MS fund has. Handy et al. (2018) describes the issues, both practical and statistical, of conducting analysis with class level data. That article analyzes class level net expense ratios, prompting them to conduct their initial analysis at the class level. However, Handy et al., also introduce a robustness technique, whereby they examine data at the fund level by identifying all classes of a MS fund by hand. As they discuss, some variables are representative of the class, for example, commission structure (front-end load or redemption fee) and class level net assets, while others, such as *MS*, board quality rating, and turnover are unique to the fund, that is, is the same for all classes.

We analyze data at the fund level, using the Stata Collapse command to create a single observation for each unique fund in the sample. The fund is the appropriate unit of analysis because governance ratings are for a fund, not for individual classes, regardless of whether the fund is MS or Non-MS. While using the fund level data ignores subtlety that class level characteristics bring to the analysis, we are interested in fund governance. As such, the traditional sample of class level observations is reduced from over 8,000 to approximately 2,300.

4.2. Empirical model

While Morningstar's Stewardship ratings are discrete rankings ranging from 1 to 5, the actual shift from one level to another is unobservable. As such, we use an ordered probit model to examine the relationship among fund characteristics and governance ratings. The model takes the following form:

$$y^*_i = X_i\beta + e_i, \text{ where } e_i \sim n(0,1). \tag{1}$$

'y^*_i' takes on the values 1 to 5, corresponding to the ordinal ranking values for *Board Qual Rate* and *Manager Incent Rate* separately. 'X_i' is a vector of independent control variables. All regression models include year-fixed effects, and standard errors are robust to heteroscedasticity.

The primary variable of interest is *MS* to test Hypothesis 1 (a, b), but *Corp Culture Rate* is also included as an independent variable of interest to test Hypotheses 2 (a, b). When *Manager Incent Rate* is the dependent variable, we include *Board Qual Rate* as an additional variable of interest to test Hypothesis 3.

Analyzing Morningstar component ratings is new; therefore, so is the empirical model. The choice of independent variables reflects possibly predictable relationships between the variables and ratings. If Morningstar's evaluation process is perfectly efficient, then we would have no a priori expectation that fund characteristics are related to ratings. However, Morningstar's process is partially judgment based, likely introducing measurement error. As such, we include variables in the model that may be correlated with governance ratings.

There are 11 common control variables across the ratings' models. *Agg Growth*, *Growth*, and *Small Cap* identify funds that are in Morningstar's aggressive growth, growth, and small cap investment objectives, respectively. Growth and income funds are the omitted category. Each variable is a dummy equal to 1 if the fund is in the respective category and 0 otherwise. *Instl* identifies funds attracting institutional investors and is a dummy variable equal to 1 if the fund, or a class in the fund, is targeted to institutional investors and 0 otherwise. *Load* identifies funds attracting retail investors in the advisor-sold channel and is a dummy variable equal to 1 if the fund, or at least one class of the fund, has a front-end load, a contingent deferred sales charge, or a level load commission structure, and 0 otherwise. *12b-1* identifies funds charging a 12b-1 fee and is a dummy equal to 1 if the fund, or at least one class of a fund, has a 12b-1 fee, and 0 otherwise. 12b-1 fees have become a primary form of compensation in advisor-sold funds, and as such, we identify this characteristic separately from *Load*. *Fund Assets* and *Family Assets* are included to capture size at the fund and fund family level, respectively. They are measured as assets under management and transformed as the natural logarithm. *Fund Age* is the age of the oldest class in the fund, and manager tenure (*Mgr Tenure*) is the longest recorded manager tenure of a class in the fund. Each variable is log transformed. Both variables, before log transformation, are measured in years. *Year* is included to control for trends in the data and takes the values 2005–2009 for each year a fund appears in the sample. Finally, when *Manager Incent Rate* is the dependent variable, we control for the fund's net expense ratio (*Netexpense*), measured as the average expense ratio across all classes of MS funds.[6]

5. Primary empirical results

5.1. Summary statistics

Summary statistics for the sample are presented in Table 1. There are over 2,300 observations across the years meeting the data analysis requirements. In column (1), approximately 73% of observations are MS funds. The results suggest that ratings are higher for Non-MS funds, with *Board Qual Rate* and *Corp Culture Rate* statistically significant at the 1% level. The average ratings for MS funds versus Non-MS funds, respectively, are: *Board Qual Rate* 3.646 versus 3.926; *Manager Incent Rate* 3.272 versus 3.322; and *Corp Culture Rate* 3.520 versus 4.206.

In the full sample, approximately 60% of funds have loads (column 1), dominated by the MS structure, where 80% of funds (column 2) have at least one load class. This is expected given that MS funds are targeted to the advisor-sold channel. However, in today's market, funds using the MS structure also include institutional classes. Sample wide, 51% of funds have an institutional representation; however, this is driven by the MS subsample, where 68% of funds have an institutional class, while only 2.6% of Non-MS funds are for institutional investors. This is evidence the MS structure has broadened since its introduction. Approximately 50% of funds have a 12b-1 fee, but again, this is driven by the MS subsample where 64% of funds have a class with a 12b-1 fee, further evidence the MS structure targets the advisor-sold channel.

Table 1 Summary statistics

Variables	Full sample			Multiple share class funds			Non-MS funds (MS = 0)			Difference	Significance
	N	(1) Mean	Standard deviation	*N*	(2) Mean	Standard deviation	*N*	(3) Mean	Standard deviation		
Board Qual Rate	2,349	3.719	0.755	1,724	3.644	0.713	625	3.926	0.825	−0.282	***
Manager Incent Rate	2,349	3.285	1.116	1,724	3.272	1.109	625	3.322	1.135	−0.049	
Corp Culture Rate	2,349	3.698	0.999	1,724	3.513	0.965	625	4.206	0.913	−0.693	***
Age in years	2,349	18.640	15.901	1,724	20.119	16.937	625	14.560	11.688	5.559	***
Fund Assets in mill	2,348	4.429	1.189	1,724	4.530	1.124	625	4.151	11.181	0.379	
Family Assets in mill	2,349	111.263	169.974	1,723	90.923	148.013	625	167.337	209.564	−76.415	***
Load	2,349	0.603	0.489	1,724	0.803	0.398	625	0.051	0.221	0.752	***
Instl	2,349	0.130	0.165	1,724	0.168	0.150	625	0.026	0.158	0.143	***
12b-1	2,349	0.303	0.248	1,724	0.401	0.210	625	0.033	0.100	0.368	***
MS	2,349	0.734	0.442	1,724	1.000	0.000	625	0.000	0.000		
Netexpense	2,349	0.940	0.369	1,724	0.945	0.346	625	0.926	0.424	0.019	
Mgr tenure in years	2,336	5.938	4.643	1,722	5.861	4.434	614	6.154	5.184	−0.293	

Note: ***$p < 0.01$.

This table presents the summary statistics for the entire sample of fund level observations in the Growth and Income, Growth, Agg Growth, and Small Cap investment objectives. We present the pooled sample results and the Non-MS and MS sub-samples for comparison. *Board Qual Rate*, *Manager Incent Rate*, and *Corp Culture Rate* are measures of board quality, managerial incentives, and fund sponsor corporate culture as evaluated by Morningstar. Each ranges from 1 (lowest) to 5 (highest). MS is equal to 1 if the fund is an MS fund and 0 otherwise. Load is a dummy equal to one if the fund has a commission structure (FEL, CDSC, or level load) associated with it and 0 otherwise. Instl equals 1 if the fund is for institutional investors or has a class for institutional investors and 0 otherwise. Fund Assets and Family Assets are measured in millions of dollars and measure the size of the fund and fund sponsor, respectively. Fund Age and Mgr Tenure are measured in years and measure the age of the fund and the length of time the manager has been with the fund. 12b-1 equals one if the fund or a class of the fund has a 12b-1 fee and 0 otherwise. Netexpense is the difference between the funds gross expense ratio and any 12b-1 fee. Year is the year in which the fund is in the sample. The final column represents the results of a *t*-test between the Non-MS and MS sub-samples for the variables. The *t*-value and the *p*-value are presented. The samples are not assumed to have equal variances. Asterisks represent significance at the 10% (*), 5% (**), and 1% (***) level, respectively.

The full sample fund family average assets under management is $111.7 billion, but this differs across fund structure. The average size for Non-MS funds (column 3) is $167.3 billion, while for the MS subsample it is $90.9 billion. The average fund has approximately $4.4 billion dollars under management, roughly equivalent across fund structure. The average fund age in the sample is 18.6 years, with MS funds having an age of 20 years versus 14.6 for Non-MS funds. The difference across structure is not surprising. From the industry's beginnings until approximately 1980, all funds were sold with a load. Many of these funds converted to the MS structure upon its approval in the 1990s. The average manager tenure is approximately 5.9 years, similar across fund structure. Finally, the average net expense ratio, defined as the fund's gross expense ratio minus the 12b-1 fee, is 94 basis points and is similar across fund structure.

5.2. Multi-variate analysis

The univariate results suggest fund governance quality may be related to fund structure. We now examine the determinants of governance ratings in a multivariate framework using

Table 2 Board Qual Rate dependent variable

Variable	(1)	(2)
Corp Culture Rate		0.417*** (0.0306)
MS	−0.333*** (0.0845)	−0.352*** (0.0794)
Load	0.132 (0.100)	0.286*** (0.0946)
Instl	−0.0403 (0.0555)	0.0222 (0.0540)
Family Assets	−0.0576*** (0.0147)	−0.0883*** (0.0148)
Family Age	−0.149*** (0.0324)	−0.0651** (0.0322)
Lnassets	0.125*** (0.0197)	0.0851*** (0.0187)
12b-1	−0.265* (0.149)	−0.0582 (0.145)
Agg Growth	−0.133 (0.110)	−0.0117 (0.114)
Growth	0.00986 (0.0540)	0.0148 (0.0541)
Small Cap	0.100 (0.0749)	0.0527 (0.0759)
Mgr Tenure	0.131*** (0.0240)	0.0807*** (0.0245)
Year	−0.0263 (0.0194)	−0.0340* (0.0193)
Year-fixed effects	Yes	Yes
Observations	2,324	2,324

Note: Robust standard errors in parentheses.
$*p < 0.1$, $**p < 0.05$, $***p < 0.01$.
This table presents ordered probit results of estimating equation (1) with *Board Qual Rate* as the dependent variable. All variables are defined as in the Table 1 heading, except *Family Assets*, *Fund Assets*, *Fund Age*, and *Mgr Tenure*, which are the log transformed value of family assets, fund net assets, fund age, and fund manager tenure. Year-fixed effects are included and standard errors are robust to heteroscedasticity. Asterisks represent significance at the 10% (*), 5% (**), and 1% (***) level, respectively.

the ordered probit model described by Equation (1). The results are presented for each dependent variable separately.

5.2.1. Board Qual Rate

Table 2 presents the results analyzing the relationship between *Board Qual Rate* and the fund characteristics discussed previously. Our primary interest is on the variables *MS* and *Corp Culture Rate*. Based on Hypothesis 1(a), we expect the coefficient estimate for *MS* to be positive and statistically significant. The results in column (1) do not support the hypothesis but do corroborate the univariate analysis above. *MS* is *negative and statistically significant*. The results indicate that MS funds are less likely (likelier) to have higher (lower) board quality scores than Non-MS funds. This result is concerning given the SEC's focus on protecting individual investors, as investors in MS funds are likely to be most vulnerable to the principal-agent problem with fund management. Table 3, Panels A and B provide predicted probabilities.[7] Panel A indicates that MS funds are three times more likely to have the lowest board quality rating (a probability of 0.4% vs. 0.1%, respectively) and are consistently more likely to have below average board quality ratings. Moreover, MS funds are significantly less likely to achieve the highest board quality rating (12.2% vs. 20%, respectively).

In column (2) of Table 2, we test Hypothesis 2(a), that is, whether fund sponsors with better corporate culture, as rated by Morningstar, are more likely to operate funds with better board quality. There is strong support for Hypothesis 2(a). *Corp Culture Rate* is positive and

Table 3 Board Qual Rate predictive probabilities

Panel A:

MS	*Board Qual Rate*				
	1	2	3	4	5
0	0.001	0.017	0.277	0.505	0.2
1	0.004	0.036	0.375	0.464	0.122
Difference (basis points)	30	190	980	−410	−780
% Difference	300.00%	111.76%	35.38%	−8.12%	−39.00%

Panel B:

		Board Qual Rate				
Corp Culture Rate	*MS*	1	2	3	4	5
1	0	0.017	0.112	0.57	0.277	0.023
1	1	0.038	0.178	0.591	0.183	0.01
	Difference (basis points)	210	660	210	−940	−130
	% Difference	123.53%	58.93%	3.68%	−33.94%	−56.52%
2	0	0.006	0.057	0.485	0.396	0.056
2	1	0.015	0.102	0.561	0.295	0.027
	Difference (basis points)	90	450	760	−1010	−290
	% Difference	150.00%	78.95%	15.67%	−25.51%	−51.79%
3	0	0.002	0.025	0.361	0.493	0.119
3	1	0.005	0.05	0.467	0.414	0.064
	Difference (basis points)	30	250	1060	−790	−550
	% Difference	150.00%	100.00%	29.36%	−16.02%	−46.22%
4	0	0	0.009	0.235	0.536	0.219
4	1	0.001	0.021	0.341	0.504	0.132
	Difference (basis points)	10	120	1060	−320	−870
	% Difference	N/A	133.33%	45.11%	−5.97%	−39.73%
5	0	0	0.003	0.134	0.507	0.356
5	1	0	0.008	0.218	0.536	0.238
	Difference	0	50	840	290	−1180
	% Difference	N/A	166.67%	62.69%	5.72%	−33.15%

This table presents predicted probabilities resulting from ordered probit regressions. *Board Qual Rate* takes on an integer value between 1 and 5 reflecting the Morningstar Board Quality Rating. *MS* equals zero if the fund is a Non-MS fund and 1 if the fund is an MS fund. The values within the table, unless labelled otherwise, represent percentages in decimal form. Panel A provides predicted probabilities based on fund type alone. Panel B provides predicted probabilities for MS and Non-MS funds conditional on their Morningstar Corporate Culture rating.

significant at the 1% level. Equally important is that our primary variable of interest, *MS*, continues to be negative and significant.

Panel B of Table 3 further details the results. Across all corporate culture ratings, MS funds are more likely to have lower board quality ratings compared with Non-MS funds and in most cases are half as likely to achieve the highest board quality ratings (a rating of 4 or 5). Interestingly, as corporate culture ratings increase, MS and Non-MS funds are less likely to have poor board quality ratings and are more likely to have higher ratings, so fund sponsor culture clearly influences governance at the fund level. The general relationship continues to hold: MS funds are more likely to have lower board quality ratings compared with Non-MS Funds. When comparing funds with the worst corporate culture ratings (a score of 1), the model predicts MS funds will have the worst board quality rating (a score of 1) with a 3.8% probability. This is double the 1.7% probability associated with Non-MS funds. For the same corporate culture rating, MS funds have only a 1% chance of receiving the highest board quality rating (a score of 5) while Non-MS funds have a 2.3% chance. Focusing on sponsors

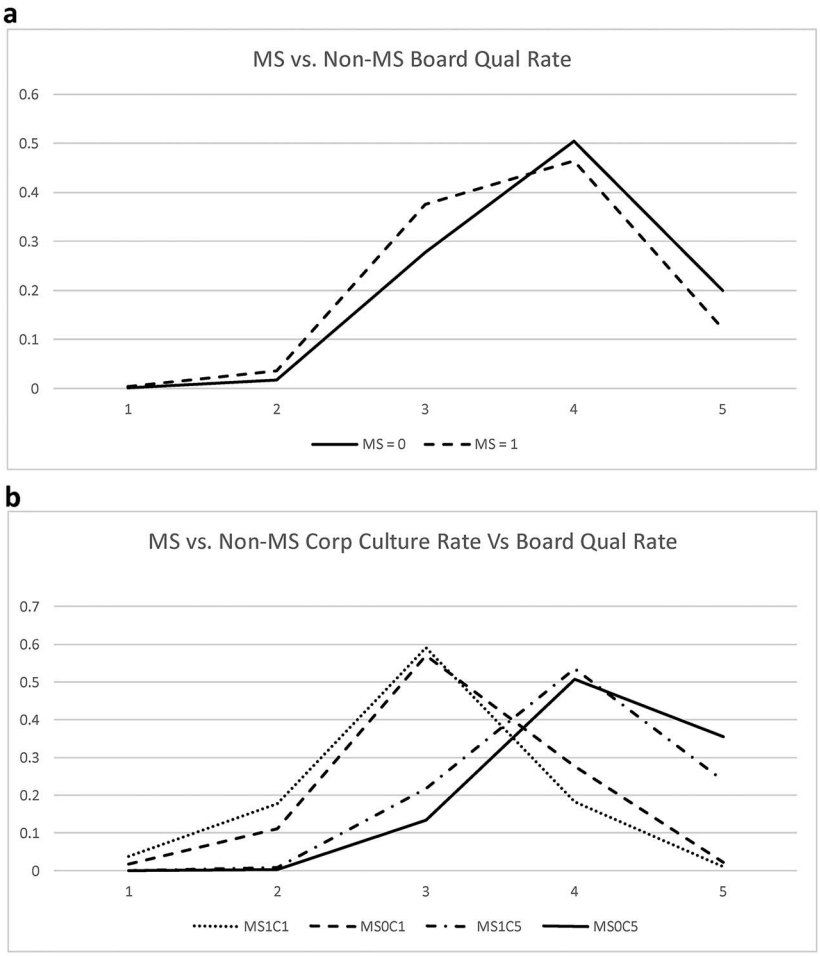

Fig. 1. MS versus Non-MS Board Qual Rate Fig. 1 compares Board Qual Rate across fund type (MS vs. non-MS; Fig. 1a) and across both fund type and Corp Culture Rate (Fig. 1b) Series *MS* = 0 reflects non-MS funds in Fig. 1a. Series *MS* = 1 reflects MS funds in Fig. 1a. Series MS1C1 graphs predicted probabilities for MS Funds with the lowest *Corp Culture Rate* (*Corp Culture Rate* = 1; Fig. 1b). Series MS1C5 graphs predicted probabilities for MS Funds with the highest *Corp Culture Rate* (*Corp Culture Rate* = 5; Fig. 1b). Series MS0C1 graphs predicted probabilities for non-MS Funds with the lowest *Corp Culture Rate* (*Corp Culture Rate* = 1; Fig. 1b). Series MS0C5 graphs predicted probabilities for non-MS Funds with the highest *Corp Culture Rate* (*Corp Culture Rate* = 5; Fig. 1b).

with only the highest corporate culture rating (a score of 5), no firms have a board quality score of 1; however, MS funds are twice as likely to receive the next worse rating (0.8% vs. 0.3%, respectively) and are significantly less likely to achieve the highest board quality score (23.8% vs. 35.6%).

A visual representation of predicted probabilities is useful. Fig. 1 provides a graphical comparison of *Board Qual Rate* across mutual fund types (MS vs. non-MS, Fig. 1a) and across mutual fund type and *Corp Culture Rate* (Fig. 1b).[8] Both figures show that non-MS funds have lower probabilities of receiving lower board quality ratings and higher

probabilities of higher board quality ratings (MS0C1 and MS0C5). At the most extreme, Fig. 1b shows that non-MS funds with the higher corporate culture ratings have approximately a 40% higher chance of receiving a top board quality rating than MS funds with the lowest corporate culture rating.

We next discuss the results for the control variables in Table 2, although for space considerations we do not provide or discuss predicted probabilities. The variables *Agg Growth*, *Growth*, and *Small Cap* are not statistically significant. However, other fund characteristics are significantly correlated with board quality. Funds with 12b-1 fees (*12b-1*) have significantly lower board quality ratings than those without a 12b-1 fee in column (1). However, when *Corp Culture Rate* is included in column (2), *12b-1* is no longer statistically significant. In contrast, *Load* is positive but not significant in column (1), but when *Corp Culture Rate* is included in column (2), *Load* is positive and statistically significant.

Looking at characteristics related to family/fund size, age, and manager tenure tells a mixed story. Larger fund families (*Family Assets*) are associated with lower *Board Qual Rate*, independent of the effect of corporate culture, as are older funds (*Fund Age*). Both results are consistent with fund families (older funds) being complacent, possibly because of prior success. Also, larger fund families may be more likely to use "captured boards," that is, board members sit on multiple boards within a family, receiving significant levels of compensation from the family.

In contrast to family size and fund age, larger funds (*Fund Assets*) and funds with longer tenured managers (*Mgr Tenure*) have higher measures of *Board Qual Rate*. The finding for *Fund Assets* may be indicative of larger funds being in the spotlight and responding to implicit pressure to provide strong governance. The finding for longer tenured managers is consistent with these managers finding value in strong governance. The final control, *Year*, indicates that on average *Board Qual Rate* is declining over the sample period, although only significantly so when *Corp Culture Rate* in included in the model.

5.2.2. Manager incent rate

Results from estimating Equation (1) with *Manager Incent Rate* as the dependent variable are presented in Table 4. As discussed above, we augment the control variables by including *Corp Culture Rate* (Hypothesis 2(b)) and *Board Qual Rate* (Hypothesis 3) in the model. Additionally, we include *Netexpense* as an additional control variable. When analyzing Hypothesis 1(b), column (1) in Table 4 shows that *MS* is negative and significant, indicating MS funds are less (more) likely to have higher (lower) managerial incentive ratings than Non-MS funds, which does not support Hypothesis 1(b).

Table 5, Panels A and B provide predicted probabilities of the variables of interest, similar to Table 3. Panel A focuses strictly on the fund structure's relationship to the managerial incentive rating. MS funds are twice as likely to have the lowest managerial incentive rating (7.3% vs. 3.5%, respectively) and are significantly less likely to achieve the highest managerial incentive rating (15.4% vs. 25.6%). In total, the findings suggest that managerial incentives are less likely aligned with retail investors in the advisor-sold channel, giving rise to a more pronounced principal-agent problem between shareholders and fund management.

Table 4 Manager Incent Rate dependent variable

Variable	(1)	(2)	(3)	(4)
Board Qual Rate		0.140*** (0.0314)		0.0955*** (0.0324)
Corporate Cult Rate			0.158*** (0.0285)	0.134*** (0.0296)
MS	−0.411*** (0.0830)	−0.385*** (0.0830)	−0.416*** (0.0842)	−0.398*** (0.0844)
Load	0.149 (0.0983)	0.133 (0.0984)	0.198* (0.102)	0.179* (0.102)
Instl	−0.144** (0.0573)	−0.139** (0.0573)	−0.118** (0.0575)	−0.118** (0.0575)
Netexpense	−0.0163 (0.0821)	0.0215 (0.0823)	0.0243 (0.0823)	0.0437 (0.0827)
Family Assets	−0.113*** (0.0164)	−0.105*** (0.0164)	−0.121*** (0.0165)	−0.115*** (0.0167)
Fund Assets	0.152*** (0.0172)	0.143*** (0.0171)	0.138*** (0.0173)	0.134*** (0.0173)
12b-1	0.450*** (0.121)	0.475*** (0.118)	0.542*** (0.122)	0.546*** (0.121)
Mgr Tenure	0.182*** (0.0240)	0.171*** (0.0242)	0.163*** (0.0241)	0.158*** (0.0242)
Fund Age	−0.0956*** (0.0293)	−0.0817*** (0.0296)	−0.0607** (0.0302)	−0.0565* (0.0303)
Agg Growth	0.114 (0.112)	0.117 (0.112)	0.153 (0.112)	0.149 (0.112)
Growth	0.0752 (0.0545)	0.0681 (0.0550)	0.0707 (0.0553)	0.0666 (0.0555)
Small Cap	0.110 (0.0753)	0.0895 (0.0762)	0.0796 (0.0765)	0.0701 (0.0769)
Year	0.133*** (0.0185)	0.136*** (0.0185)	0.132*** (0.0183)	0.134*** (0.0184)
Year-Fixed Effects	Yes	Yes	Yes	Yes
Observations	2,324	2,324	2,324	2,324

Note: Robust standard errors in parentheses.

$*p < 0.1$, $**p < 0.05$, $***p < 0.01$.

This table presents ordered probit results of estimating Equation (1) with *Manager Incent Rate* as the dependent variable. All variables are defined as in the Table 1 heading, except *Family Assets*, *Fund Assets*, *Fund Age*, and *Mgr Tenure*, which are the log transformed value of family assets, fund net assets, fund age, and fund manager tenure. Year-fixed effects are included and standard errors are robust to heteroscedasticity. Asterisks represent significance at the 10% (*), 5% (**), and 1% (***) level, respectively.

As argued above, corporate culture and board quality may influence managerial incentive quality independently. The results in columns (2–4) of Table 4 provide support for the hypotheses that higher corporate culture ratings lead to higher managerial incentive ratings (Hypothesis 2(b)) and higher board quality ratings lead to higher managerial incentive ratings (Hypothesis 3), independently, and when both are included in the model (column (4)). Funds with higher *Corp Culture Rate* and higher *Board Qual Rate* are more likely to have higher managerial incentive ratings, a sign of strong sponsor and fund level governance. MS funds continue to have lower managerial incentive ratings. The significance of the *MS* variable in Table 4 when compared with the univariate findings in Table 1 highlights the importance of multivariate analysis.

Panel B of Table 5 presents the analysis of how the predicted probability of achieving a certain managerial incentive rating changes across fund type and either, the corporate culture rating or the board quality rating. The results are similar across rating type and echo results in Panel B of Table 3. MS funds with the lowest corporate culture ratings are likelier to have the lowest managerial incentive ratings with a 12.9% probability (compared with Non-MS' 6.9% probability). Additionally, MS funds with the lowest corporate culture ratings are also significantly less likely to have the highest managerial incentive ratings (8.8% vs. 16.2%, respectively). As corporate culture improves, both MS and Non-MS funds are more likely to achieve higher managerial incentive ratings; however, MS funds are more likely to achieve lower scores and less likely to achieve higher scores compared with Non-MS funds. For example, funds having a corporate culture rating of 5, Non-MS funds have a 31% probability of achieving the highest managerial incentive rating, while MS funds only have a 19.5%

Table 5 Manager Incent Rate Predictive Probabilities Man*ager Incent Rate* takes on an integer value between 1 and 5 reflecting the Morningstar

Panel A:

	Manager Incent Rate					
	MS	1	2	3	4	5
	0	0.035	0.119	0.342	0.248	0.256
	1	0.073	0.182	0.384	0.209	0.154
Difference (basis points)		380	630	420	−390	−1020
% Difference	108.57%	52.94%	12.28%	−15.73%	−39.84%	

Panel B:

	Manager Incent Rate							*Manager Incent Rate*					
Corp Culture Rate	*MS*	1	2	3	4	5	*Board Qual Rate*	*MS*	1	2	3	4	5
1	0	0.069	0.175	0.381	0.213	0.162	1	0	0.056	0.159	0.376	0.225	0.184
1	1	0.129	0.24	0.383	0.16	0.088	1	1	0.11	0.225	0.389	0.174	0.103
	Diff. (BPs)	600	650	20	−530	−740		Diff. (BPs)	540	660	130	−510	−810
	% Diff	86.96%	37.14%	0.52%	−24.88%	−45.68%		%Diff	96.43%	41.51%	3.46%	−22.67%	−44.02%
2	0	0.054	0.153	0.37	0.228	0.194	2	0	0.047	0.143	0.366	0.235	0.209
2	1	0.106	0.219	0.388	0.179	0.109	2	1	0.094	0.209	0.39	0.187	0.119
	Diff. (BPs)	520	660	180	−490	−850		Diff. (BPs)	470	660	240	−480	−900
	%Diff	96.30%	43.14%	4.86%	−21.49%	−43.81%		%Diff	100.00%	46.15%	6.56%	−20.43%	−43.06%
3	0	0.042	0.133	0.355	0.241	0.229	3	0	0.039	0.128	0.354	0.244	0.235
3	1	0.085	0.197	0.387	0.197	0.134	3	1	0.081	0.193	0.388	0.2	0.138
	Diff. (BPs)	430	640	320	−440	−950		Diff. (BPs)	420	650	340	−440	−970
	%Diff	102.38%	48.12%	9.01%	−18.26%	−41.48%		%Diff	107.69%	50.78%	9.60%	−18.03%	−41.28%
4	0	0.032	0.113	0.336	0.251	0.268	4	0	0.032	0.114	0.34	0.251	0.263
4	1	0.068	0.175	0.381	0.214	0.162	4	1	0.068	0.177	0.384	0.212	0.158
	Diff. (BPs)	360	620	450	−370	−1060		Diff. (BPs)	360	630	440	−390	−1050
	%Diff	112.50%	54.87%	13.39%	−14.74%	−39.55%	%Diff	N/A	55.26%	12.94%	−15.54%	−39.92%	
5	0	0.025	0.095	0.313	0.257	0.31	5	0	0.027	0.101	0.324	0.256	0.292
5	1	0.054	0.153	0.37	0.229	0.195	5	1	0.058	0.161	0.377	0.224	0.18
	Diff. (BPs)	290	580	570	−280	−1150		Diff. (BPs)	310	600	530	−320	−1120
	%Diff	116.00%	61.05%	18.21%	−10.89%	−37.10%		%Diff	114.81%	59.41%	16.36%	−12.50%	−38.36%

Managerial Incentive rating. *MS* equals zero if the fund is a Non-MS fund and 1 if the fund is an MS fund. The values within the table, unless labelled otherwise, represent percentages in decimal form. Panel A provides predicted probabilities based on fund type alone. Panel B provides predicted probabilities for MS and Non-MS funds conditional on their Morningstar Corporate Culture rating or the Morningstar Board Quality rating.

probability. The board quality results echo the corporate culture findings. In short, governance matters and appears stronger in Non-MS funds, as rated by Morningstar.

The results for the control variables capturing load status, fund objectives, family size, fund age, fund size, and manager tenure are all similar to those in Table 2. In all columns of Table 4, *Instl* is negative and statistically significant, indicating funds having an institutional presence have lower managerial incentive ratings. In all columns, the presence of a 12b-1 fee leads to increased *Manager Incent Rate*. Finally, the inclusion of *Netexpense* to the models adds no explanatory power. Finally, unlike in Table 2, *Year* is positive and significant, indicating that on average managerial incentive ratings are increasing during the sample period.

6. Conclusion

While mutual funds have been widely studied, the bulk of the work has been in the area of performance and cost structure. There has been much less work, theoretical or empirical, examining the MS fund structure introduced widely in 1995. The evidence that does exist is not complimentary of the structure.

The analysis in this article finds further evidence highlighting problems with the MS structure. MS funds have board quality and managerial incentive ratings that would suggest that they, and the companies sponsoring them, tend to align themselves with investor interests less. Specifically, the governance quality appears lower than in Non-MS funds as determined by an independent rating source, Morningstar. While we do not attribute causality to these results, they do identify another unfavorable characteristic associated with the MS structure. The results remain consistent as we account for changes in other governance variables.

The findings in this article suggest that in addition to the risk of being directed toward classes with suboptimal costs, as suggested by O'Neal (1999), advisor-led retail investors are also, simply by investing in MS funds, investing in funds whose governance is subpar compared with Non-MS funds. This is concerning as investors in the advisor-sold distribution channel have been shown to be less financially savvy. In total, evidence is mounting against the MS structure.

Notes

1 We take the Morningstar ratings as given and assume that they are unbiased and useful as tools to evaluate board quality and managerial incentives, that is, this article does not evaluate board or managerial effectiveness directly.

2 Here, the term "costs" encompasses commissions *and* ongoing fund expenses. Often, the term cost is used interchangeably with "ongoing fund expenses" or the "expense ratio."

3 It is important to note that the authors do not mean the fund's investment advisor when referencing "advisor" but rather mean the investor's financial advisor (broker, wealth manager etc.).

4 We recognize that a financial advisor could also fulfill some, or all, of this oversight role, although literature cited above suggests that they may not.

5 In an effort to update the sample used in this article, the authors contacted Morningstar sales and research departments and were advised that Morningstar stopped providing stewardship ratings in approximately 2017. Additionally, Morningstar indicated that historical data was no longer available.
6 See Handy et al. (2018) for how fund expenses and managerial incentives may be related.
7 The results remain consistent when evaluating predicted probabilities holding all other variables at the mean here and in the remainder of the article.
8 We provide the graphical representation for the results in Table 2 only. To conserve space, we do not provide figures for Table 4, although the conclusions mirror those in Figure 1 and are available upon request.

References

Alexander, G. J., Jones, J. D., & Nigro, P. J. (1998). Mutual fund shareholders: Characteristics, investor knowledge, and sources of information. *Financial Services Review, 7,* 301–316.

Barber, B. M., Odean, T., & Zheng, L. (2005). Out of sight, out of mind: The effects of expenses on mutual fund flows. *Journal of Business, 78,* 2095–2120.

Bergstresser, D., Chalmers, J. M. R., & Tufano, P. (2009). Assessing the costs and benefits of brokers in the mutual fund industry. *The Review of Financial Studies, 22,* 4129–4156.

Cao, J., Ghosh, A., Goh, J., & Ng, S. (2014). *Governance Matter: Morningstar Stewardship Grades and Mutual Fund Performance.* Working Paper, Singapore Management University.

Capon, N., Fitzsimons, G. J., & Prince, R. A. (1996). An individual level analysis of the mutual fund investment decision. *Journal of Financial Services Research, 10,* 59–82.

Chou, J., Ng, L., & Wang, Q. (2011). Are better governed funds better monitors? *Journal of Corporate Finance, 17,* 1254–1271.

Fact Sheet. (2006). *The Morningstar Stewardship Grade for Funds.* (available at https://quicktake.morningstar.com/DataDefs/StewGradeMethodology.pdf).

Handy, J., Nichols, H., & Smythe, T. (2018). Should investors care about mutual fund governance quality?: Evidence from Morningstar stewardship ratings. *The Journal of Wealth Management, 21,* 44–58.

Jones, M. A., Lesseig, V. P., & Smythe, T. I. (2005a). Financial advisors and mutual fund selection. *Journal of Financial Planning, 18,* 64–70.

Jones, M. A., Lesseig, V. P., & Smythe, T. I. (2005b). Financial advisors and multiple share class mutual funds. *Financial Services Review, 14,* 1–20.

Lesseig, V. P., Long, D. M., & Smythe, T. I. (2002). Gains to mutual fund sponsors offering multiple share class funds. *Journal of Financial Research, 25,* 81–98.

Livingston, M., & O'Neal, E. S. (1998). The cost of mutual fund distribution fees. *Journal of Financial Research, 21,* 205–218.

Moore, S., & Porter, G. E. (2017). Can investors benefit from using Morningstar's stewardship grades? *Journal of Accounting and Finance, 17,* 130.

Nanda, V. K., Wang, Z. J., & Zheng, L. (2009). The ABCs of mutual funds: On the introduction of multiple share classes. *Journal of Financial Intermediation, 18,* 329–361.

Nofsinger, J. R., & Varma, A. (2007). How analytical is your financial advisor? *Financial Services Review, 16,* 245–260.

O'Neal, E. S. (1999). Mutual fund share classes and broker incentives. *Financial Analysts Journal, September-October,* 76–87.

Securities and Exchange Commission Rule 18f-3 Adoption Release, Release No. 33-7143, IC-20915, 1995.

Financial Services Review announces

Call for Papers for two Special Issues on:

- **Financial literacy**
- **The Value of Financial Advice**

Guest editor:

Terrance K Martin Jr.
Woodbury School of Business
Utah Valley University

Submission deadline: March 31, 2020

Financial Services Review (FSR) is the official publication of the Academy of Financial Services (AFS). The primary objective of this refereed academic journal is to encourage rigorous empirical research that examines individual behavior in terms of financial planning and services. We are pleased to announce a call for papers for two special issues on "**Financial Literacy**" and "**The Value of Financial Advice**." Original research papers relevant to each of the topics are welcomed. The empirical implications of submitted manuscripts should be relevant for academic and practitioners in the field of individual financial management.

Our goal is to attract papers that will increase scientific knowledge in these two important areas of research in financial planning and individual financial management. Therefore, we encourage new perspectives and application or extension of previously utilized models and techniques. Unfocused, poorly written papers, or papers without data and empirical analysis will not be considered. All submissions will be double-blind peer-reviewed.

Papers submitted to the Special Issues will receive expedited reviews. We are also offering a reduced submission/membership fee of $100 (total) for submission to these special issues. In your submission, **please indicate which Special Issue you are submitting**.

Submission guidelines for the Financial Services Review can be found at this link https://academyfinancial.org/Submission-Guidelines and https://academyfinancial.org/special-issues

Papers should be e-mailed to the FSR Editor, Stuart Michelson, smichels@stetson.edu and guest Editor, Terrance K Martin Jr., Terrance.martin@uvu.edu, Submission fees (at the reduced rate of $100) may be paid at https://academyfinancial.org/special-issues

For further questions or inquiries about the special issue, please contact the guest editor, Terrance K Martin Jr., Terrance.martin@uvu.edu or Financial Services Review Editor, Stuart Michelson, smichels@stetson.edu.

CE 1 hour general principles of financial planning/investment planning

AFS and FPA members can earn CE credits through Financial Services Review. Go to FPAJournal.org.

To receive one hour of continuing education credit allotted for this exam, you must answer four out of five questions correctly. CFP Board recently adopted revisions to several provisions of its CE policies, including changing the minimum number of questions for self-study assessments from 10 to 5 per full CE credit hour. Therefore, Financial Services Review CE exams will have 5 questions. CE credit for this issue expires October 31, 2021, subject to any changes dictated by the CFP Board. AFS and FPA offer Financial Services Review CE online only -–- paper continuing education will not be processed. Go to FPAJournal.org to take current and past CE (free to AFS and FPA members). You may use this page for reference. Please allow 2-3 weeks for credit to be processed and reported to CFP Board.

1. In "Are 'Fun' Sources of Windfalls Destined to be Spent Hedonistically?" by Bland and Chambers, when receiving a windfall of unexpected cash, how do people generally use it?
    a. As rationally as possible.
    b. Emotionally, usually without reason or logic.
    c. Mostly rationally, but it depends in part on the source of the income.
    d. Mostly on large, durable assets.

2. In Bland and Chambers, when receiving a windfall, how much do people just spend on fun?
    a. All of it, indicating one can't have too much fun.
    b. None of it, indicating that we're serious adults.
    c. A fixed percent of it, indicating that we stick to our flexible budgets.
    d. A large chunk of a small refund, tapering off for big refunds, indicating fun in moderation.

3. In Bland and Chambers, how can financial planners use this research in combination with other academic literature?
    a. Early increased communication, especially regarding planning for the next yearly tax refunds, could be beneficial to clients and planners.
    b. Financial planners should consistently poll clients about whether they have recently received a sizeable windfall.
    c. Clients make spontaneous decisions, so approach clients shortly after tax season about refunds just received.
    d. Clients need time to process information, so approach clients several months after tax season about refunds most recently received.

4. In "A Portfolio of Leveraged Exchange Traded Funds" by Trainor, Chhachhi, & Brown, leveraged exchange traded funds magnify the daily return of an underlying index. What is the main cause for the realized leverage ratio to generally decline over time?
    a. Implied interest cost
    b. Volatility of returns
    c. Tracking error
    d. Fund expenses

5. In Trainor, Chhachhi, & Brown, investing one third of an investor's wealth into a 3x leveraged exchange traded (LETF) is equivalent to 100% invested in the underlying index. Why will this strategy not work for the typical buy-and-hold investor?
    a. The need for frequent rebalancing to maintain a set exposure.
    b. LETFs are riskier and appropriate only for short-term investors.
    c. LETFs are leveraged and require margin accounts to trade.
    d. Over the long-run, LETF's values tend to decline due to the constant daily leverage ratio.

MANUSCRIPT SUBMISSIONS and STYLE

(1) Papers must be in English.

(2) Papers for publication should be sent to the Editor: Professor Stuart Michelson, E-mail: smichels@stetson.edu. Electronic (Email) submission of manuscripts is encouraged, and procedures are discussed below.

There is a $100 submission fee payable to the Academy of Financial Services (AFS) if at least one of the authors is a member of AFS. Submission fees should be paid online at academy financial org. If none of the authors is a member of AFS, please complete an online membership application form, which can be downloaded at http://academyfinancial.org, and pay online ($225 total; $125 for a one-year membership and $100 submission fee). Submission of a paper will be held to imply that it contains original unpublished work and is not being considered for publication elsewhere. The Editor does not accept responsibility for damage or loss of papers submitted. Upon acceptance of an article, author(s) transfer copyright of the article to the Academy of Financial Services. This transfer will ensure the widest possible dissemination.

(3) Submission of papers: Authors should submit their papers electronically as an e-mail attachment to the Editor at smichels@stetson.edu. Please send the paper in Word format. Do not sent PDFs. Ensure that the letter 'l' and digit '1', and also the letter 'O' and digit '0' are used properly, and format your article (tabs, indents, etc.) consistently. Do not allow your word processor to introduce word breaks and do not use a justified layout. Please adhere strictly to the general instructions below on style, arrangement and, in particular, the reference style of the journal.

(4) Manuscripts should be double spaced, with one-inch margins, and printed on one side of the paper only. All pages should be numbered consecutively, starting with the title page. Titles and subtitles should be short. References, tables, and legends for the figures should be printed on separate pages.

(5) The first page of the manuscript, the Title Page, must contain the following information: (i) the title; (ii) the name(s), title, institutional affiliation(s), address, telephone number, fax number and e-mail addresses of all the author(s) with a clear indication of which is the corresponding author; (iii) at least one classification code according to the Classification System for Journal Articles as used by the Journal of Economic Literature, which can be found at http://www.aeaweb.org/journal/elclasjn.html; in addition, up to five key words should be supplied.

(6) Information on grants received can be given in a footnote on the Title page.

(7) The abstract, consisting of no more than 100 words, should appear alone on page 2, titled, Abstract.

(8) Footnotes should be kept to a minimum and should only contain material that is not essential to the understanding of the article. As a rule of thumb, have one or less footnote, on average, per two pages of text.

(9) Displayed formulae should be numbered consecutively throughout the manuscript as (1), (2), etc. against the right-hand margin of the page. In cases where the derivation of formulae has been abbreviated, it is of great help to the referees if the full derivation can be presented on a separate sheet (not to be published).

(10) The Financial Services Review journal (FSR) follows the APA Publication Manual, 6th Edition, style. However, consistent with the current trend followed by other publications in the area of finance, the journal has a very strong preference for articles that are written in the present tense throughout.

References to publications should be as follows: ''Smith (1992) reports that'' or ''This problem has been studied previously (Ho, Milevsky, & Robinson, 1999).'' The author should make sure that there is a strict one-to-one correspondence between the names and years in the text and those on the reference list.

The list of references should appear at the end of the main text (after any appendices, but before tables and legends for figures). It should be double spaced and listed in alphabetical order by author's name. References should appear as follows:

Books:

Hawawini, G. & Swary, I. (1990). Mergers and acquisitions in the U.S. banking industry: Evidence from the capital markets. Amsterdam: North Holland.

Chapter in a book:

Brunner, K. & Meltzer, A. H. (1990). Money supply. In: B. M. Friedman & F. H. Hahn (Eds.), *Handbook of monetary economics* (Vol. 1, pp. 357-396). Amsterdam: North Holland.

Periodicals:

Ang, J. S. & Fatemi, A. M. (1997). Personal bankruptcy costs: their relevance and some estimates. *Financial Services Review, 6,* 77-96.

Note that journal titles should not be abbreviated.

(11) Illustrations will be reproduced photographically from originals supplied by the author; they will not be redrawn by the publisher. Please provide all illustrations in quadruplicate (one high-contrast original and three photocopies). Care should be taken that lettering and symbols are of a comparable size. The illustrations should not be inserted in the text, and should be marked on the back with figure number, title of paper, and author's name. All graphs and diagrams should be referred to as figures, and should be numbered consecutively in the text in Arabic numerals. Illustration for papers submitted as electronic manuscripts should be in traditional form. The journal is not printed in color, so all graphs and illustrations should be in black and white.

(12) Tables should be numbered consecutively in the text in Arabic numerals and printed on separate sheets.

Any manuscript which does not conform to the above instructions will be returned for the necessary revision before publication.

Page proofs will be sent to the corresponding author. Proofs should be corrected carefully; the responsibility for detecting errors lies with the author. Corrections should be restricted to instances in which the proof is at variance with the manuscript. Extensive alterations will be charged. Reprints of your article are available at cost if they are ordered when the proof is returned.

Financial Services Review

(ISSN: 1057-0810)

Academy of Financial Services
Stuart Michelson
Stetson University
School of Business
421 N. Woodland Blvd.
Unit 8398
DeLand, FL 32723

(Address Service Requested)