

Volume 23

Issue 1

Spring 2025

Editor:

Olgun Fuat Sahin, Saint Louis University

Associate Editors:

Larry Bauer, Memorial University of Newfoundland
Seongsu Kim, California State University, Dominguez Hills
Won Yong Kim, Augsburg University

Editorial Board:

Sunghan Bae, Truman State University
Keshav R. Bhattarai, University of Mississippi
Tim Carpenter, Roanoke College
Leo H. Chan, Utah Valley University
Jiun-Lin (Alex) Chen, Valparaiso University
Paul Choi, Howard University
Anna N. Danielova, McMaster University
Ding Ding, The Australian National University
Thomas A. Hanson, Butler University

Yan He, Indiana University Southeast
Seth Hoelscher, Missouri State University
Jin-Gil Jeong, Howard University
Sunghoon Joo, California State University, Dominguez Hills
Sang Baum Kang, Illinois Institute of Technology
Srinidhi Kanuri, University of Southern Mississippi
Dongyoung Kim, California State University San Marcos
CNV Krishnan, Case Western Reserve University
Gisung Moon, Columbus State University
Erin Oldford, University of Regina
Wenjing Ouyang, University of the Pacific
Nischala Reddy, University of Central Missouri
Pattarake Sarajoti, Chulalongkorn University
Amit Sinha, Bradley University
Chonawee Supatgiat, Chulalongkorn University
Hui-Ju Tsai, Washington College
Yongdong Wang, Bowling Green State University

Letter From the Editor

Olgun Fuat Sahin

Original Articles

Erroneous Ecumenicism: The folly of faith fund fungibility

Michael R. Williams

Macroeconomic Determinants of the Credit Loss Forecasting

Zilong Liu Hongyan Liang Chang Liu Yang Cheng

Corporate Governance and Cash Holdings Decisions by U.S. Life Insurance Companies

Raja Bouzouita

Wherefore Art Thou?: Small business, healthcare coverage, and group association

Roger Lee Mendoza

Small Business Administration Lending and Economic Performance

Chu-Sheng Tai

Table of Contents

i Letter From the Editor
Olgun Fuat Sahin

Original Articles

- 1 Erroneous Ecumenicism: The folly of faith fund fungibility**
Michael R. Williams
- 36 Macroeconomic Determinants of the Credit Loss Forecasting**
Zilong Liu, Hongyan Liang, Chang Liu, and Yang Cheng
- 59 Corporate Governance and Cash Holdings Decisions by U.S. Life Insurance Companies**
Raja Bouzouita
- 72 Wherefore Art Thou?: Small business, healthcare coverage, and group association**
Roger Lee Mendoza
- 97 Small Business Administration Lending and Economic Performance**
Chu-Sheng Tai

To the Members of the Academy of Finance and Readers of the Journal of Finance Issues,

Welcome to the latest issue (Volume 23, Number 1) of the *Journal of Finance Issues*. We are excited to present a collection of timely and impactful research that addresses various facets of the financial world. Our commitment remains to publish high-quality work that advances both academic understanding and practical application in the finance discipline.

This issue features five insightful papers:

First, in "**Erroneous Ecumenicism: The folly of faith fund fungibility**" by **Michael R. Williams**, the author investigates whether faith-based investment funds, specifically those associated with Roman Catholicism and Protestantism, demonstrate systematic differences. The paper finds that these ostensibly similar faith foci exhibit distinct risk and return characteristics, particularly in how they relate to common asset pricing factors and have become increasingly distinct over time. Williams argues that religion-specific asset pricing factors exist, cautioning against treating faith-based funds as interchangeable investment tools.

Next, "**Macroeconomic Determinants of the Credit Loss Forecasting**" by **Zilong Liu, Hongyan Liang, and Chang Liu**, addresses the critical need for robust credit loss forecasting models, especially in light of unprecedented economic volatility like the COVID-19 pandemic. The study examines the dynamic relationship between charge-off rates for U.S. commercial banks' loan portfolios (commercial and industrial, consumer, and real estate loans) and macroeconomic factors. Tai identifies key macroeconomic variables, such as labor market indicators, housing market dynamics, and consumer financial conditions, as drivers of credit risk across these loan categories, highlighting the need to re-evaluate traditional relationships for model robustness in the current economic environment.

Following that, "**Corporate Governance and Cash Holdings Decisions by U.S. Life Insurance Companies**" by **Raja Bouzouita**, examines the influence of corporate governance on cash holdings within the U.S. life insurance industry. Using a dynamic panel model, the paper demonstrates that corporate governance proxies, such as board size and independent directors, are significant factors explaining cash holding decisions and effectively mitigate the principal-agent conflict. The analysis also reveals that the adjustment rate to target cash holdings varies across U.S. life insurance companies with the corporate governance proxies used.

We also present "**Wherefore Art Thou? Small business, healthcare coverage, and group association**" by **Roger Lee Mendoza**, the author explores why small businesses manage to offer employee healthcare coverage despite substantial costs and statutory exemptions. The study investigates strategic differences between plans offered by small and large groups, focusing on how small firms can mitigate costs and risks through plan design, financing, risk management, and associational options like Association Health Plans (AHPs). Mendoza's work highlights the importance of minimizing transaction costs and efficient communication in achieving cost-efficiency in employer healthcare coverage for small groups.

Finally, in "**Small Business Administration Lending and Economic Performance**" by **Chu-Sheng Tai**. This paper analyzes state-level data from 1990 to 2023 to determine the impact of U.S. Small Business Administration (SBA) lending through its 504 and 7(a) loan programs on state-level real GDP and unemployment rates. The study finds that SBA lending generally promotes state-level real GDP and reduces unemployment rates, with statistically significant and economically important effects. However, this positive effect diminished during the 2007-2009 global financial crisis.

I extend my sincere gratitude to all the anonymous reviewers whose rigorous and insightful feedback is indispensable to maintaining the quality of our publications. Their dedication ensures the scholarly integrity and relevance of each article.

I would also like to express my deepest appreciation to our Associate Editors, Larry, David, and Won. Their tireless efforts, expertise, and commitment to the peer review process are vital to the smooth and effective operation of the journal. Their contributions are invaluable in bringing high-quality research to our readership.

We hope you find this issue informative and thought-provoking.

Sincerely,

Olgun Fuat Sahin
Editor, Journal of Finance Issues
jfi-slu@slu.edu

Erroneous Ecumenicism: The folly of faith fund fungibility

Michael R. Williams*

Abstract

This paper examines two often-associated religious beliefs, Roman Catholicism and Protestantism, and whether their respective faith based investment funds systematically differ. We find that Catholic- and Protestant- oriented investment funds exhibit different risk and return characteristics, especially in how they relate to common asset pricing factors. Also, these superficially similar faith foci have become increasingly distinct over time and show only a moderate relationship with Islamic- and Israeli- focused investment indexes. Further, we provide strong evidence of the existence of different, religion-specific asset pricing factors that are strongly (weakly) associated with more similar (dissimilar) religious funds. An extensive examination of the extant literature and these beliefs' underlying theology augment these empirical findings. Together, our findings clearly show that different faith based funds from different faiths exhibit different risk and return characteristics and that these different funds are suitable only to particular investor clienteles. This study concludes by providing implications and recommendations to academics for future research and investment participants regarding best practices; that all should avoid the easy call of investment related erroneous ecumenicism.

Keywords: Socially Responsible Investing, Faith Based Investing, factor models, time-varying correlations, faith, belief, Roman Catholicism, Islam, Jewish, Protestantism, non-financial utility, sin stocks, saint stocks, investment management, portfolio management, asset pricing models.

JEL Codes: C32, G11, G12, G18, G23, Z11

I. INTRODUCTION

The universe of Socially Responsible Investing (SRI) literature is indeed a vast and ever growing pool of knowledge, synthesizing an array of economic, financial, econometric, psychological, historical, social, and political fields into a combined body. Adding to the size and complexity of the SRI literature is that the very concept of what constitutes SRI is equally large. The SRI umbrella covers a broad investment landscape of partially-overlapping interests such as environment, sustainability, governance, and even denomination-specific religious investments. This paper examines a particular aspect of the SRI universe, Faith-Based Investing (FBI), and a largely unaddressed topic that potentially and subtly biases both researchers and market participants, alike: erroneous ecumenicism.

Specifically, we examine the four dominant FBI categories, Roman Catholicism, the more Biblically-literal Protestantism, and for contrast the Islamic and Jewish faith traditions, to examine whether these faiths should be treated as equivalents within the FBI and SRI space or whether they should be treated as very distinct categories of investment interest. If these faiths are indeed equivalent, there are implications to investment managers and investors, alike. For

* Michael R. Williams, mwilliams15@govst.edu, Governors State University.

example, funds of various faiths may be offered to those seeking risk- and sin- diversification, relative to the overall secular market, and that employing a combination of diversely-focused FBIs is appropriate; that is: treating funds of different faiths as fungible investment tools. However, if these faiths cannot be treated as equivalents, then managers and investors should be cautioned in promoting or participating in cross-faith investing; that clearly different investment foci fulfill very different investor non-financial utility demand.

This study uses three main approaches to address two questions: "are faith funds fungible?" and "is investment ecumenicism erroneous?". First, we provide a comprehensive review of the SRI and FBI literature, especially as it relates to how FBI funds differ and how different faith-driven investors also differ. We show from the prior literature that FBI supply and FBI demand show clear distinctions as both individual groupings and in their interactions; that faith-specific FBI funds are designed to, and are successful at, fulfilling faith-specific FBI demand of the funds' clients.

Second, we perform a detailed analysis (see Literature Review and Appendix) between the two main distinctions among Christian-oriented FBI funds: Roman Catholicism and Biblically-literal Christianity of which Protestantism, in its purest form, most closely aligns. Using the backdrop of Biblical Christianity, we show that Roman Catholicism and Biblical Christianity are not easily reconciled, with very strong differences existing between the two with respect to vital concepts revolving around the mechanisms and methods of salvation, the authority given to and within their respective ecclesiastical structures, the sources and fixedness of spiritual knowledge, the religious compliance believers must maintain, and more.

As such, we make a compelling case that Roman Catholic and Biblically-literal Christians have very different views on walking their respective spiritual paths, and that these different views explain the clear, consistent differences found between their associated FBI funds in the academic literature. For additional contrast, we also compare elements of godhead, salvation, assurance, and obedience within these already-distinct Christian faiths against those of the Islamic faith. Here, we show that the Islamic tradition is vastly different from Roman Catholicism and Biblical Christianity and on grounds that simply cannot be bridged; that where notions of the godhead and of salvation are so different, ecumenical mending between Christianity and Islam is simply not tenable. As such, investment managers and investing clients are encouraged to maintain vigilance over ecumenical calls, especially in the investment space, when it comes to exclusionary screening behavior. Catholic, Biblical Christian, and Islamic followers are also cautioned of efforts made by ecclesiastical bodies, and their related investment advising groups, to blend the faiths together to common practices for the purpose of advancing potentially conflicting, human secular goals.

Third, we examine 65 Catholic- and Protestant- based funds for potential differences in their time series behavior, particularly relating to their comovement and sensitivity to various asset pricing factors. Looking only at unconditional correlations, both across the sample and in a time-varying manner, we show that Catholic and Protestant funds are fairly similar, initially lending credibility to those who treat the two religions, and their faith-specific investing clienteles, similarly. Yet, using a multi-factor model, we show that Catholic and Protestant funds respond differently to common asset pricing factors with Protestant funds generally experiencing a larger range of variation in factor sensitivities and regardless of official fund labels. Using this common factor model across the funds, we find that the conditional correlations (i.e. correlations across model-filtered returns) between Catholic and Protestant funds are largely near zero, indicating little to no relationship between "religious residuals" or "faith-specific factors".

Examining rolling, time-varying conditional correlations shows an even richer picture with different faiths comoving differently and ever-more distinctly over time.

Exploring this concept of a systematic, faith-based factor further, we conduct risk- and interfaith- factor regressions on Catholic and Protestant indexes. Here, we find that Catholic and Protestant index returns are largely determined by common asset pricing factors and that the factor loadings differ to a non-trivial degree between the faiths. Thus, even after accounting for fund-idiosyncratic factors, there are systematic outcome differences between Catholic and Protestant funds, indicating systematic differences in their management. Further, modeling Catholic index returns as a function of common asset pricing factors along with the Protestant index, shows that the influence of the common factors decrease sharply while the Protestant index coefficient is significant, both statistically and economically. Furthermore, when separately modeling both the Catholic and Protestant indexes as a function of the common risk factors and indexes of Islamic and Israeli funds, only the Israeli index is significant; statistically, but not economically. Not only do we again show that Catholic and Protestant funds are mostly distinct, but that we also show that more-distant religious traditions are treated more-distinctly by FBI investors.

Thus, confirming, clarifying, and extending the existing pool of SRI and FBI literature, we demonstrate that different faiths have different underlying beliefs, which in turn, lead to different investment management practices and, by consequence, different investment risk and return outcomes for investors. We show that the tendency by SRI and FBI participants, academic researchers in particular, towards faith ecumenicism is erroneous, that faith-specific funds are not fungible, and that different funds applying different screening criteria according to different faiths should be treated differently.

This paper continues in Section Two with a detailed examination of the SRI and FBI literature. Section Three details our empirical methodology and Section Four presents the results of our empirical findings. Section Five augments these findings with a discussion of the implications-to and recommendations-for FBI participants, including academic researchers, investment fund managers, the investors themselves, and financial regulators. Finally, an Appendix has been provided at the end which performs a comparison and contrast of the three, ecumenically-conflated religious faiths: Roman Catholicism, Biblical Christianity, and Islam.

II. ACADEMIC LITERATURE

Faith based basics

At the highest level, Socially Responsible Investing (SRI) can be thought of as allocating any portion of one's limited investment funds towards non-financial purposes (Dunfee, 2003). And, despite there being a large variation in, and overlap among, the definitions of various SRI fund labels (Junkus & Berry, 2015), SRI tends to be segmented according to green-, ethical-, social-, and religious- investing styles (Ito et al., 2013). Thus, Faith Based Investing (FBI) is a niche subset of SRI (Adams & Ahmed, 2013).

The root of both SRI and FBI is grounded in portfolio theory and its inability to completely account for or explain non-financial aspects of investing utility, such as social welfare, faith, and so on (Markowitz, 1952; Wan-Ni, 2012). As such, portfolio optimization and stock selection are not entirely driven by financial factors (Kim & Venkatachalam, 2006) as investors have multi-factor utility functions that take into account non-financial aspects. More

specifically, investors not only consider utilitarian, performance-based factors when allocating across and within their portfolios, but also according to expressive, non-financial utility drivers (Statman, 2004; Bollen, 2007). As such, while investors certainly have performance considerations within their utility optimization decisions, non-financial considerations are also at play; a significant one being faith and salvation (Hood et al., 2013).

FBI, then, is a byproduct of investors making allocation and selection decisions on non-financial, faith-based grounds. What's more, as different investors hold sometimes very different perspectives on what constitutes "sinful" or unacceptable behavior (Borgers & Pownall, 2014; Niszczoła & Bialek, 2021a, 2021b; Meunier & Ohadi, 2023), these different perspectives create investor clienteles and segmented markets (Renneboog et al., 2011; Hood et al., 2013), even along faith-based lines. Thus, FBI can be defined as integrating any spiritual, non-financial aspect into one's investing criteria, aligning one's investing behavior with spiritual beliefs (Czerwonka, 2014), and where outcomes may include the pursuit of faith-driven returns, abnormal returns, or a combination of the two (Wan-Ni, 2012).

FBI is not, by any means, a new concept. Negative screening effectively began in the 1600s with the Quakers eschewing slave-trade investments (Kinder & Domini, 1997; Louche et al., 2012), 1920s Methodists banning alcohol, defense, and gambling investments (Renneboog et al., 2008b), and the very first SRI/FBI fund (The Pioneer Fund; 1928) applying active alcohol, gambling, and tobacco screening filters to its investment decisions (Kinder & Domini, 1997; Junkus & Berry, 2015). Since those early days, FBI has been a growing trend within the United States, seeing widespread acceptance in 1971 of the Quaker- and Methodist- influenced Pax World Fund (Kreander et al., 2004), and even internationally with, for example, Christian churches being instrumental in the South African apartheid divestment movement (Teoh et al., 1999).

Negative screening sin

To address their clients' unique investing criteria and objectives, FBI managers generally use negative screening (i.e. avoiding unacceptable investments), positive screening (i.e. seeking acceptable investments), engagement strategies (i.e. interacting with less-than-acceptable investments in hopes of initiating internal change), and impact strategies (i.e. actively participating in or with an investment in hope of enacting external change). Based on their examination of 100, largely Christians institutional investors, Louche et al. (2012) find that 87% of respondents use negative screening, 79% various positive approaches, and 88% shareholder activism approaches, with the top five US negative screens focusing on nuclear weapons, defense, tobacco, abortion, and pornography. Diener and Habisch (2022b) find similar results for Christian-focused funds with 100% of surveyed funds utilizing negative screening, 78% positive approaches, and 64% shareholder activism. Interestingly enough, 70% of these Christian-oriented funds employed non-financial, "green" criteria revolving around the environment and society. Statman and Glushkov (2009) found similar negative screening filters within the broader category of SRI funds: alcohol, defense, gambling, and tobacco.

Over time, FBI funds have shifted their focus from negative screening alone towards active, impact investing (Diener, 2022). Yet, the efficacy of non-exclusionary approaches is questioned with some (e.g. Derwall et al., 2011; Slager & Chapple, 2015) seeing no positive influence on constituent firm behavior, others (e.g. Clementino & Perkins, 2020) finding an unintended negative effect, while still others show a positive influence on the firms (e.g. Diener

& Habisch, 2022a). This is not surprising in that Enete and Kiss (2018) find that about half of all firms in the Vanguard S&P 500 Index have one or more anti-Christian attributes: 88 firms violating abortion screens, 50 pornography, 26 gambling, 12 entertainment, 8 tobacco, and 1 human rights.

Negative screening by SRI and FBI managers is not without its reported downsides. Specifically, negative screening may attenuate risk-adjusted returns, depending on the assumed pricing model (Chang & Witte, 2010) and with greater screening intensity causing greater impairment (Renneboog et al., 2008b). Renneboog et al. (2008a) clarify that while raw returns are impaired by about 2.2% to 6.5% per year, risk adjusted returns are not impacted. In addition to being a potential drag on upside returns, negative screening imposes direct costs to the fund of about 0.17% to 2.40% per annum and with these costs impacting investors both immediately and cumulatively (Adler & Kritzman, 2008). Geczy et al. (2005) find that negative screening (unscreened) SRI funds have an average 1.3% (1.1%) annual expense ratio. Finally, negatively screened funds face lower diversification benefits due to these funds' restricted asset universes (Rudd, 1981; diBartolomeo & Kurtz, 1999; Barnett & Salomon, 2006).

Negative screening funds tend to have size, scalability, and sustainability issues (Davis et al., 2017). Further, and despite the 67% of surveyed Christian-focused funds having ten or more investment screens (Diener & Habisch, 2022b), negative screening funds may simply fall short of their clients' screening expectations (Schwartz, 2003). Kurtz and diBartolomeo (2005) note the difficulty in detecting firms for exclusion in screening criteria, stem cell research being a prime example. Dion (2009) finds that, when applying United States Conference of Catholic Bishops (USCCB) -influenced investment screens, at least one of the three Christian-focused funds contained almost 30% of the "controversial companies" that should have been screened out. Gellman (2013) finds that one popular Catholic-oriented fund's criteria only screens out about 5% of the firms from the Russell 3000 Index, most of which would have been screened out on performance grounds alone. This type of "screening slippage" means that fund screens and moral stances may not fully predict what a fund actually achieves.

Nevertheless, negative screening can come with positive benefits. Some find that socially screened funds (Sauer, 1997; Boasson et al., 2004; Bauer et al., 2006; Bauer et al., 2007; Cortez et al., 2009) and faith screened funds (Boasson et al., 2006; Lyn & Zychowicz, 2010) do not underperform unscreened funds. Others find that stricter exclusion screens lead to higher investment returns with, according to some authors, returns being the highest for the most extreme high- and low- levels of screening intensity (Barnett & Salomon, 2006; Lee et al., 2010). Diltz (1995) finds a more nuanced picture where screening for environmental and defense sectors leads to positive performance gains but that other screening types do not. Geczy et al. (2005) find that negative screening leads to lower fund turnover while Renneboog et al. (2011) note that employing in-house screening teams produces positive, risk-adjusted returns. Further, Liston and Soydemir (2010) find that imposing negative screens does not lead to a diversification cost or benefit.

Despite the costs negative screening may impose on clients, these clients have shown a willingness to pay (Enete & Kiss, 2018), given their moral stances. Vrecko and Langer (2013) find that investors are willing to pay between 3.31% (4.03%) across a one (five) year investment window in order to customize their returns distribution via altering their investment product. Meunier and Ohadi (2023) find that two-thirds of US investors are willing to pay 2.3% of their initial investment to specify excluded industries. Thus, despite the numerical costs of negative screening, investment managers are not necessarily short-changing clients (Boasson et al., 2004).

Faith based results

Given faith-based investors' unique, non-financial investment preference sets and some of the real-world costs of implementing FBI, we next look at how FBI performs relative to both non-faith and secular-SRI investments; not surprisingly, there is a range of research findings on this issue. Some find that FBI performance is not statistically different from secular-benchmarks (Boasson et al., 2006; Wan-Ni, 2012) or conventional funds (Davis et al., 2017). Kurtz and diBartolomeo (2005) show that the KLD Catholic Values 400 Index performs about the same as the S&P 500 Index while Naber (2001) shows that Catholic-screened funds do not underperform unscreened funds. Thus, according to some of the literature, faith-based indexes are suitable for those wanting to adhere to their faith without expecting abnormal returns (Wan-Ni, 2012); that the lack of underperformance suggests that FBI managers can apply ethical screens without sacrificing financial performance or potentially violating their fiduciary duty to their clients (Boasson et al., 2006).

Other portions of the literature contend that US-based FBI funds outperform both secular-SRI and market indexes (Lyn & Zychowicz, 2010) whereas others suggest that any FBI overperformance is a product of time variation; that negative screening can weed out unethical firms and those with poor governance, hence avoiding crash and crisis risk (Boasson et al., 2004). Still others demonstrate that so-called "sin investments" not only outperform the market as a whole, but FBI in particular; that FBI also underperforms the market (Liston & Soydemir, 2010). Carosella et al. (2012) find that the KLD Catholic Values 400 Index underperforms on a risk-adjusted basis and that lower SRI-scoring stocks have higher risk-adjusted returns. It is not surprising, then, that Ferruz et al. (2012) find FBI mutual fund managers underperform the market and secular mutual funds due to FBI managers eschewing sin stocks and not being able to time style changes within the market.

Faith based differences

Looking more broadly into the distinctions between the secular world and those of faith, we find clear differences and at multiple levels. From the corporate world, we know that greater general religiosity within a region leads to less risky firm policies (Hilary & Hui, 2009), less risk taking (Callen & Fang, 2015), less firm financial distress (Gharbi et al., 2020), better financial reporting quality (Montenegro, 2017), and less opportunism against minority stockholders (Chen et al., 2016). Christianity and Judaism, in particular, are associated with greater economic freedom and activity, entrepreneurial activity, gender and minority equality, social progress, care for the environment, lower corruption (Russell et al., 2023), and greater tax compliance (Boone et al., 2013). Further, self-identification with Christianity is positively associated with ecological, social, and ethical activities (Gutsche, 2019), including SRI participation (Czerwonka, 2014). Christian-identifying higher education institutions are relatively more involved with SRI activities, ESG concerns, and rejecting sin stocks; Catholic and Baptist institutions being the most strict (Smith & Smith, 2016).

On an individual level, religious distinctions play a definite role in risk aversion and risk taking, with higher levels of religiosity being associated with more risk aversion (Miller & Hoffmann, 1995; Diaz, 2000; Bartke & Schwarze, 2008), Catholics and Protestants, in particular (Leon & Pfeifer, 2017). Religious households tend to hold less risky assets, develop more savings, and are more thrifty (Renneboog & Spaenjers, 2012). This effect is only weakly related

to religious upbringing, yet is positively related to current religious activity (e.g. church attendance; Noussair et al., 2013). Shu et al. (2012) disagree with the above by saying that Catholics (Protestants) are actually more (less) risk-seeking than the general population. Yet, this disagreement may be due to a need to parse risk-types, given that Catholics are equally risk-adverse as the general population for *pure* risks (i.e. risks with no potential for upside) but more risk-seeking for *speculative* risks (i.e. risks with both upside and downside; Halek & Eisenhauer, 2001).

Looking particularly at religiosity and financial decisions, believers align their investing behavior with their beliefs (Czerwonka, 2014). Religiosity impacts not only attitudes on risk taking and gambling, but also non-gambling decisions (Kumar, 2009) such as portfolio choices at the individual level and asset prices at a macro level (Kumar et al., 2011; Page & Spalt, 2011). Thus, religiosity can influence financial markets via a risk aversion channel and individual gambling preferences (Hilary & Hui, 2009), can lower stock volatility (in Catholic countries; Blau, 2017), and even influence the performance of sin-related stocks (Salaber, 2013). These influences can alter the behavior of investment vehicles like hedge funds, causing those funds in strongly Catholic countries to hold less risky stocks, have better cross-industry diversification, lower total- and idiosyncratic- risk, and all of this especially during times of market crises (Gao et al., 2017).

As Leon and Pfeifer (2017) note, "*...it is necessary to control for underlying general risk-taking preferences due to the heterogeneous attitudes of Christians and Muslims regarding financial risk taking...*", meaning that specific religiosity can impact asset pricing and returns differently. But what about the differences between seemingly-similar faiths such as Catholicism and Protestantism? Areas of high Christian Protestantism are positively correlated with economic freedom (Hillman & Potrafke, 2018), creditor rights (Stulz & Williamson, 2003), and the likelihood of regional venture capital firms being risk-averse and IPO-active, than with Catholic-intense areas (Chircop et al., 2020). Evangelical Protestantism, in particular, is negatively correlated with financial statement irregularities (McGuire et al., 2012).

On the subject of different sin perspectives and relative to Catholics, Protestants generally view gambling as sinful (Ellison & Nybroten, 1999), live in areas that are less likely to adopt state lotteries and have robust lottery sales (Kumar et al., 2011), are less likely to legalize gambling (Johnson & Meier, 1990), while supporting gambling and liquor regulations (Fairbanks, 1997). Catholics, on the other hand, tend to be more permissive towards gambling, including bingo and other charitable games, and even use these activities as fundraising sources (Hoffman, 2000; Halek & Eisenhauer, 2001).

At the asset level, even though different religious faiths exhibit common, overlapping screening criteria (Wan-Ni, 2012), these overlaps are non-trivial in nature; clear distinctions exist among how inter-Christian faiths participate in asset markets. For example, Protestants, relative to Catholics, are more "sin adverse", are less likely to hold sin stocks (Hood et al., 2013; Kumar & Page, 2013), and require a greater return premium for investing in sin stocks (Salaber, 2007; Salaber, 2009). As noted above, religious differences can even impact secular, sin-related stocks where these stocks' prices are more undervalued in Protestant countries, leading to an approximate 8.5% sin stock return premium; an effect not seen in Catholic countries. This is consistent with Protestant investors' relative aversion to sin stocks and is consistent with a form of the Segmented Market Hypothesis, based on religious preferences (Salaber, 2013).

Looking specifically at risk and return preference differences, Protestants are more risk averse than Catholics (Noussair et al., 2013). Parsing risk shows that both Catholics and

Protestants are equally risk averse for *pure* risks but with Catholics being more risk seeking for *speculative* risks (Shu et al., 2012). Investors within predominantly-Catholic (relative to Protestant) regions tend to hold stocks with lottery features (i.e. firms with high idiosyncratic risk) and prefer employee stock option plans (Kumar et al., 2011; Page & Spalt, 2011). Further, mutual funds in predominantly-Catholic areas have higher fund risk taking, volatility, investment concentration, turnover, inter-reporting-period trading aggressiveness, and tournament risk taking behavior (i.e. chasing lagging first half returns with more aggressive risk taking in the second half of a given year; Shu et al., 2012).

There are also differences among religious orientations at the portfolio screening level. Specifically, Islamic funds tend to negatively screen firms focused on usury, pork, and entertainment, whereas Catholic funds tend to screen out firms dealing with adult entertainment and alternative lifestyles. In fact, across a variety of different investment screening criteria, 37.5% of these screening categories are very different between Catholic and Islamic funds; they are both, also, very different in their respective negative screening behaviors for "green" or "ethical" concerns (Ghoul & Karam, 2007).

Looking only between Catholic and Protestant funds, Diener and Habisch (2022b) note that "*...by and large, there is no 'Christian investing' in the sense of an ethical investment policy that most fund providers have similarly implemented... the policies are determined by differing approaches to interpreting Biblical texts and by divergent social and political influence factors...*". Here, they find that Catholics and Protestants differ most in their screening behavior regarding contraception, embryonic stem cell research, alcohol, defense, gambling, and environmental concerns. And, while they find that political and social factors are important drivers of screening behavior, the Biblical-interpretation motivation is the most influential factor. As is found in our analysis of Roman Catholicism and Biblically-literal Christianity (see Appendix), these findings are not surprising: Catholic funds are being driven by the use of a more-dynamic, progressive interpretation of the Bible whereas Biblically-literal funds base their screening and investing behavior on a more-fixed, conservative interpretation of the Bible.

Finally, the differences seen above can be interdenominational in nature. Specifically, Evangelical (i.e. Fundamentalist or Biblical Literalists), relative to both Catholics and (moderately-literal) Protestants, are more likely to avoid companies that deal with abortion, alcohol, gambling, homosexual lifestyles, and pornography (Statman, 2005). Also, some Christian denominations (e.g. Lutherans) are adverse to altering fund objectives based on changing, secular social norms (e.g. "social justice"; Inskeep, 1992).

Erroneous ecumenical license

Yet, despite the prior literature consistently highlighting clear-cut differences in interreligious- and interdenominational- screening propensities, risk tolerances, risk outcomes, and more, the existing literature is still prone to erroneous ecumenicism in its use of terms like "faith", "religion", "Christian", and "Bible"; glossing over real-world differences among very different religious beliefs which, in turn, may gloss over important differences among different faiths' investment behavior and outcomes. This is not to say that this ecumenicism is intentional or even malicious to any degree, but that as academic researchers and industry practitioners alike, we should always strive to strain out nuanced gnats from our results.

One of the ways that studies ecumenically gloss over important differences among religions, and even denominations, is to speak of different religions, implicitly or explicitly, as

being one. An example of this is when religiosity is homogenized to include any type of religious practice, regardless of potentially differing underlying beliefs (Miller & Hoffmann, 1995; Chen et al., 2016). Another example is when religious designations are simply left ambiguous (Gharbi et al., 2020). A common way this appears in the literature is for different faiths-centric investment indexes or funds to be aggregated into a single index, and all religions within implicitly treated as one (e.g. Areal et al., 2010; Liston & Soydemir, 2010; Lyn & Zychowicz, 2010; Ferruz et al., 2012; Nofsinger & Varma, 2014; Davis et al., 2017).

Alternatively, some studies will treat one faith, for example Roman Catholicism, as being representative of all of Christianity (e.g. Ghoul & Karam, 2007; Czerwonka, 2014; Harjoto & Rossi, 2019; Rossi et al., 2019). Naturally, some Protestants and Biblical Literalists would suggest that Roman Catholicism's focus on works-based salvation, claims made regarding Jesus' mother Mary, the use of non-inspired additions to the previously-closed canon of the Bible, and so on, may disqualify Catholicism as "Christianity's" unitary representative. Still other literature provides excellent, religiously-disaggregated results that show non-trivial differences between religious investing behavior (e.g. negative screening criteria), and yet go on to gloss over their clearly different results with suggestions of marginal differences (e.g. Ghoul & Karam, 2007). Other authors knowingly homogenize different religious distinctions when presenting or summarizing their results, even though their empirical analyses contain disaggregated results with material differences in investing behavior or outcomes (e.g. Boasson et al., 2004; Bartke & Schwarze, 2008; Chen & Diaz, 2012; Beer et al., 2014; Gao et al., 2017; Mao et al., 2024).

More subtle than the above, is when academic studies understandably look for and to common FBI management standards, but where these standards, and standards-generating bodies, are unwarrantedly ecumenical in their efforts. Examples of some of these groups and efforts include: Socially Responsible Investment Guidelines for the United States Conference of Catholic Bishops (USCCB), Investing Guidelines of the (Catholic) Austrian Bishop's Conference, Catholic Papal Encyclicals, guidelines produced by various Protestant denomination and Christian-adjacent groups, and even the human-secularist UN Sustainable Development Goals (SDGs) and Ecumenical Office of the United Nations. The Catholic-oriented Interfaith Center on Corporate Responsibility (ICCR), as another example, includes such far-flung Christ-centric traditions as Benedictine-, Dominican-, Jesuit-, and Marian- Catholicism, the Episcopalian Church, Presbyterians, Evangelical Lutherans, Methodists, and even Christ-optional Unitarian Universalists. Aiding and assisting the ICCR's efforts is its Laudes- and Open Society- Foundation supported Investor Alliance for Human Rights, which brings together more than 240 investment firm titans, representing 21 countries, and over \$14 trillion in assets; creating a common platform for coordinating and communicating industry standards, even to some of the largest and most well-recognized Christ-centric investment companies (investorsforhumanrights.org/about).

These coordinating bodies introduce, both implicitly and explicitly, shared investment standards that have the potential to overshadow FBI managers' faith-based efforts with largely human secular endeavors revolving around themes like the climate, world governance, cross-national migration, race and culture, economic egalitarianism, corporate and societal transformation, and more; sometimes with these efforts being couched, however loosely, in and with religious undertones. In net, ecumenical pressures are being exerted by and within the implicit nexus of ecclesial bodies, supranational entities, private foundations, investment companies, and even academic research.

Underlying faith differences

Naturally, with such clear differences being seen in the prior literature regarding different faith-focused funds being managed differently, demanded differently, and empirically responding differently, it warrants a review of whether these differences are justifiably attributed to faith differences or some other factor. Should different religions be considered and kept as distinct, or simply merged into a type of ecumenically blended set of faith "best practices"?

To start, various efforts over many centuries have attempted to lump the so-called "Abrahamic faith traditions" into ecumenical equivalency. This includes, recently, the various United Nations endeavors (e.g. the Ecumenical Office of the United Nations) and the Higher Committee of Human Fraternity's Abrahamic Faith House, both of which are working to bring the Catholic, Islamic, and Jewish traditions together under a common umbrella. Here, ecumenicism is seen by non-fundamentalists as a chance to bridge seemingly minor doctrinal gaps that, once spanned, will allow the faiths to work together in terms of peace, prosperity, and other typically human-centric objectives.

Yet, even a cursory examination of the core, distinguishing features among Biblical Christianity, Roman Catholicism, and Islam reveal stark contrasts in how these three faiths handle critical issues of both deity and salvation. We provide a detailed analysis of these differences in the Appendix of this paper. We encourage readers unsure of the differences among these faiths, or the concerns of erroneous ecumenicism, to examine the Appendix for a nuanced discussion. For the purposes of a summary: the Appendix makes clear that Biblically-literal Christianity, Roman Catholicism, and Islam are three, very different faiths with stark differences in their views of God's attributes, the foundational source-material for their faith, the depravity of man, the mechanisms of salvation necessary to resolve issues of sin, and how one continues to maintain obedience to God. These differences, in turn, shape how investment participants view risk and return tradeoffs, faith-consistent investment behavior, and more. Consequently, we should not be surprised if the investing behavior of faith-based managers, and the investment demand of faith-based investors, should differ markedly among the oftentimes erroneously homogenized faiths. As a foreshadowing: our empirical results show exactly that.

III. METHODOLOGY

Sample construction

Our sample of FBI funds primarily includes the two, main seemingly-related Christ-centric faiths: Roman Catholic-aligned funds and Biblically-literal Protestant funds. By examining the extant literature and performing various internet searches, we find a total of 150 Christ-oriented FBI funds, 19 of which are Exchange Traded Funds and 131 are mutual funds. Where possible, the institutional class of each mutual fund is employed. From this initial sample, we delete funds that are classified as real estate, hedge funds (e.g. strong derivatives exposure, long-short funds, etc.), target date funds, and fixed income strategies. Additionally, we choose a minimum fund inception date of January 3rd, 2017; balancing the desire for as many funds as possible, while ensuring a robust sample size and stability of trading volume. This deletion process leaves us with 65 funds: 12 Catholic-focused and 53 Protestant-focused. The table below reports various characteristic statistics of the remaining funds:

Table 1: Sample Statistics

The following table reports the number and percentage (of total) of each type of Chris-focused fund represented in the sample. The sample representations are for both the universe of identified, publicly-traded, Christ-focused investment funds without (Pre-Truncate) and with (Post-Truncate) fund elimination for investment strategy (i.e. Real Estate, Hedge Fund, Fixed Income, and Target Date) and inception date (i.e. funds with inception dates before January 3rd, 2017) filters. Differences in the various samples are provided for illustration. Additional classifications are made according to funds' investment overall strategy (Type), size preference (Size), detailed strategy (Strat), location preference (Loc), and religious focus (Rel). All investment funds noted in the Post-Truncate, Catholic, and Protestant samples have data spanning on or before January 2nd, 2018 to the end of the sample of December 31st, 2024.

		Pre-Truncate		Post-Truncate		Diff	Catholic		Protestant		Diff
Type	Blend	76	62.3%	39	60.0%	-2.3%	6	50.0%	33	62.3%	-12.3%
	Growth	23	18.9%	16	24.6%	5.8%	5	41.7%	11	20.8%	20.9%
	Value	23	18.9%	10	15.4%	-3.5%	1	8.3%	9	17.0%	-8.6%
Size	Large	73	59.8%	43	66.2%	6.3%	7	58.3%	36	67.9%	-9.6%
	Medium	34	27.9%	12	18.5%	-9.4%	2	16.7%	10	18.9%	-2.2%
	Small	15	12.3%	10	15.4%	3.1%	3	25.0%	7	13.2%	11.8%
Strat	Divid	3	3.0%	2	3.1%	0.1%	1	8.3%	1	1.9%	6.4%
	Equity	86	85.1%	52	80.0%	-5.1%	11	91.7%	41	77.4%	14.3%
	Mixed	12	11.9%	11	16.9%	5.0%	0	0.0%	11	20.8%	-20.8%
	Real Est	2									
	Hedge	8									
	Target	11									
	Fix Inc	28									
Loc	US	129	86.0%	53	81.5%	-4.5%	10	83.3%	43	81.1%	2.2%
	World	21	14.0%	12	18.5%	4.5%	2	16.7%	10	18.9%	-2.2%
Rel	Catholic	35	23.3%	12	18.5%	-4.9%					
	Protest	115	76.7%	53	81.5%	4.9%					

From Table 1, we note that the Pre- and Post- Truncation samples are very similar, indicating that the filtering of funds for strategy- and date- characteristics does not artificially bias the fund Type (i.e. Blend, Growth, and Value), Size (i.e. Large, Medium, and Small), Strategy (i.e. Dividend, Equity, and Mixed), Location (i.e. US and World), or Religious (i.e. Catholic and Protestant) foci. Additionally, when looking across the Post-Truncation sample and across Religious classifications, we see that both the Catholic and Protestant samples are mostly balanced among the various fund classification types. Granted, the Catholic funds tend to be relatively more focused on Growth Styles and Smaller Sizes, yet the samples are still mostly matched. This finding means that any future results, especially those regarding factor loadings or fund classifications, will not be biased by our sample construction method. Ultimately, the sample's daily price data, collected from Yahoo!Finance spans January 2nd, 2018 to December 31st, 2024; allowing late-inception funds (i.e. those that started near 2017) at least a full year to mature and stabilize trading.

Factor and Filtering Regressions

Throughout this study, we examine the various FBI funds using both raw- and filtered- returns as well as the estimated factor loadings from the following regression:

$$r_{i,t} = \alpha_i + \beta_{1,i}MKT_t + \beta_{2,i}SMB_t + \beta_{3,i}HML_t + \beta_{4,i}INV_t + \beta_{5,i}MOM_t + \beta_{6,i}WLD_t + \varepsilon_{i,t} \quad (1)$$

which combines elements from the Fama French (1995) Five Factor Model and the Carhart (1997) Four Factor Model (mainly: the Momentum factor). Here, factors are selected to capture the likely investment style characteristics of each fund and allow filtered returns (i.e. estimated regression residuals) to be free of non-Religious influences; the estimated residuals will, hence, be approximate measures of a given FBI fund's religious dynamics.

Also, for Equitation 1 and beyond, *MKT* proxies for systematic market risk (i.e. SPDR S&P 500 ETF returns), *SMB* proxies for the small minus big factor (i.e. Vanguard Small-Cap ETF returns minus Vanguard Mega Cap ETF returns), *HML* proxies for the high book-to-market (i.e. value) minus low book-to-market (i.e. growth) factor (i.e. Vanguard Value ETF returns minus Vanguard Growth ETF returns), *INV* proxies for the conservative minus aggressive factor (i.e. iShares Core Conservative Allocation ETF returns minus iShares Core Aggressive Allocation ETF returns), *MOM* proxies for the momentum factor (i.e. iShares MSCI USA Momentum Factor ETF returns), and *WLD* proxies for an international investment factor (i.e. iShares Core MSCI Total International Stock ETF returns). Note that we use ETF-based returns to proxy for the various risk factors for two reasons. First, at the time of writing, the factor returns generously made available on the Kenneth R. French Data Library did not extend to the full sample length of our study. Second, using ETF-based risk factors allows our factors to be explicitly market-based as opposed to the more implicitly-based Fama French Factors.

Faith Indexes and comovements

In addition to the above Factor-only model, we also estimate OLS regressions for religious-centric models. Specifically, we estimate the following OLS regression for an equally-weighted index of Catholic ($r_{C,t}$) and Protestant (variously $Prot_t$ and $r_{P,t}$) fund returns:

$$r_{C,t} = \alpha_0 + \beta_{1,C}MKT_t + \beta_{2,C}SMB_t + \beta_{3,C}HML_t + \beta_{4,C}INV_t + \beta_{5,C}MOM_t + \beta_{6,C}WLD_t + \lambda_{1,C}Prot_t + \varepsilon_{C,t} \quad (2)$$

We also estimate the following OLS regression on Catholic ($r_{C,t}$) and Protestant ($r_{P,t}$) fund returns using equally-weighted indexes of Islamic and Israeli fund returns:

$$r_{i,t} = \alpha_i + \beta_{1,i}MKT_t + \beta_{2,i}SMB_t + \beta_{3,i}HML_t + \beta_{4,i}INV_t + \beta_{5,i}MOM_t + \beta_{6,i}WLD_t + \lambda_{1,i}Islam_t + \lambda_{2,i}Israel_t + \varepsilon_{i,t} \quad (3)$$

and where $Islam_t$ is an equally-weighted measure of Islamic-centric fund returns (i.e. Wahed FTSE USA Shariah ETF, SP Funds Dow Jones Global Sukuk ETF, and SP Funds S&P 500 Sharia Industry Exclusions ETF) and $Israel_t$ is an equally-weighted measure of Israeli-centric fund returns (i.e. VanEck Israel ETF and ARK Israel Innovative Technology ETF). Note that the primary focus of this study is Catholic and Protestant-centric funds whereas a limited number of Islamic- and Israeli- investment funds are incorporated into the analysis in the form of faith-specific indexes (see below) and for the limited purpose of contrasting the Christ-centric funds

with non-Christ-centric funds. The results from the Equation 2 and 3 estimations are not intended to be exhaustive, but only illustrative. The rationale for this more limited, non-Christ-centric fund approach is this: if empirical differences can be detected between Catholic and Protestant funds representing ostensibly-similar faiths, then equal or larger differences should be found outside of these two faiths. Additionally, taking a more limited approach to the non-Christ-centric funds limits the dimensionality of the analysis. Future research is encouraged to tackle this issue more comprehensively.

IV. EMPIRICAL RESULTS

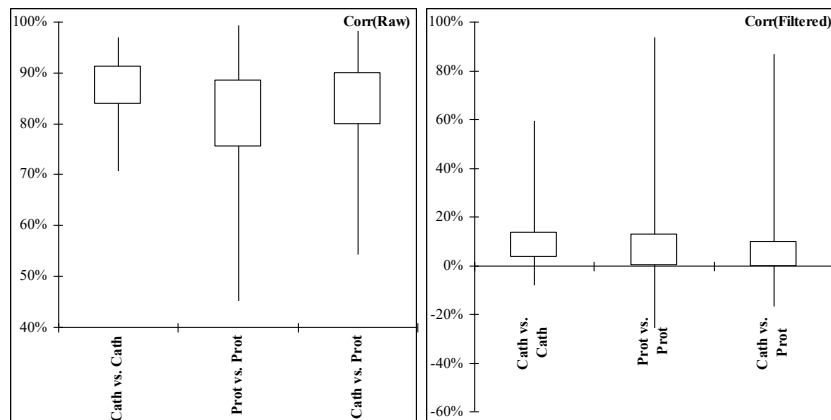
Having shown that Roman Catholicism and the more Biblically-literal Protestantism have very different characteristics in the existing literature (Section Two) and in their respective religious beliefs (see Appendix), we now examine whether such differences can be systematically detected through correlation and factor sensitivity analyses.

Correlation Box Plots

The Figure 1 box plots below report the minimum, 25th percentile, 75th percentile, and maximum pair-wise correlations among Catholic- and Protestant- centric funds. Examining the left box plot of raw return correlations, we see that all fund pairs are highly correlated regardless of whether the fund pair is Catholic/Catholic, Protestant/Protestant, or Catholic/Protestant. We find that the range of variability in Protestant/Protestant pair correlations is higher than for the Catholic/Catholic and Catholic/Protestant pairs. This theme of Protestant fund variability will be seen across most of the subsequent results.

Figure 1: Fund Correlations

The following box plots display various correlation coefficients calculated on both the raw and filtered returns of 65 Christ-focused funds (12 Catholic and 53 Protestant) across the entire sample (January 2nd, 2018 to December 31st, 2024; n = 1,761). Filtered returns are the estimated residuals from a full-sample, OLS estimation of Equation 1. Each box plot reports the maximum, 75th percentile, 25th percentile, and minimum correlation coefficient values with the left box plot depicting the raw return correlations and the right box plot depicting the filtered return correlations:



In general, the raw return correlations indicate that the different faith fund classifications are superficially correlated, suggesting that the different faith funds are not different in outcomes. This also initially appears to suggest that differences between the Catholic and Protestant faiths are not reflected in the funds based on these religions.

Yet, when examining the right box plot of factor-filtered return correlations, we find that the own- and cross- faith correlations drop to near zero. With the exception of outlier correlations, Catholic/Catholic and Protestant/Protestant fund pair correlations are weakly positive and the Catholic/Protestant pair grouping less so. Thus, filtering for common fund style, size, and other investment factors indicates that Catholic- and Protestant- centric funds are substantially different; that different faith funds are achieving different market outcomes due to them being managed in very different ways. Thus, we have preliminary evidence that different faiths are being reflected in their respect funds in the form of demonstrably different outcomes.

Regression Box Plots

Figure 2 presents the results of various OLS estimations of Equation 1 for the 65 faith-based funds. For the coefficient box plots below, we separate the results into three main classifications: all funds (All), Catholic funds (Cath), and Protestant funds (Prot). Additional fund- and style-classification groupings were examined in unreported results, providing consistent evidence as the aggregated categories presented here. Detailed results from the factor regressions are available upon request.

From the Adjusted- R^2 results in Figure 2, we see that the multi-factor model (Equation 1) does an exemplary job at modeling and explaining faith fund returns. While there are outlier estimations, Adjusted- R^2 s north of 0.85 are the norm. As seen in the correlation box plot results and most elsewhere, the Protestant funds have the highest range of variability in terms of both the 25th/75th percentile range and maximum/minimum range. While the differences detected in the Adjusted- R^2 results are not as stark as the correlation box plot results or subsequent estimation results, the theme remains: Catholic and Protestant funds are exhibiting non-negligible different market outcomes. Also, given the greater range of Adjusted- R^2 variability for the Protestant-centric funds, Protestant funds may be driven by faith-specific influences more so than the factor-explained Catholic funds.

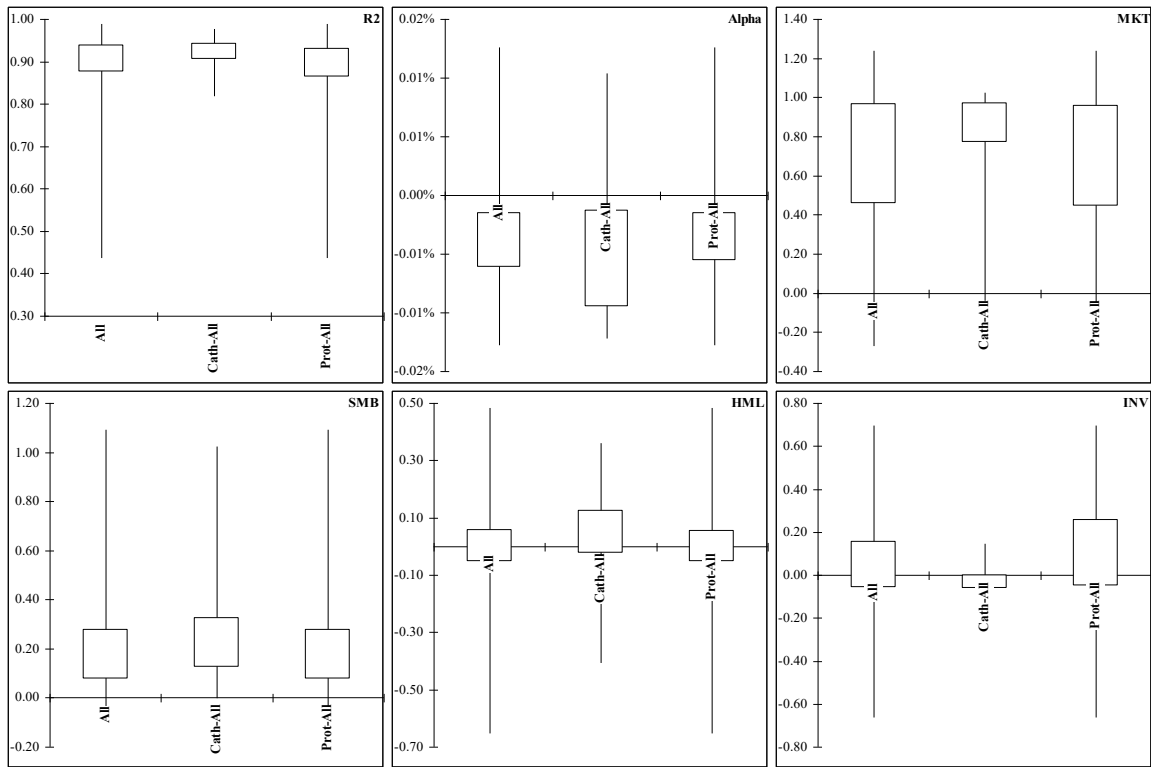
Examining the estimated Alphas, we find in Figure 2 that, with the exception of outlier results, most faith funds have negative alphas indicating that most faith fund managers are not producing superior risk-adjusted performance nor are they adept at timing the market. Unlike the Adjusted- R^2 results, the differences in Catholic and Protestant funds is more pronounced. For the *MKT* factor loading, we find that Catholic- and Protestant- focused funds are highly related to overall, systematic market risk and that, in general, Protestant fund factor loadings are much more varied than those of Catholic funds. Again, this suggests that Protestant-related funds are being driven by other risk factors as well as idiosyncratic, religion-specific aspects. Put differently: Protestant-centric investment funds experience different asset pricing characteristics, likely due to their underlying screening- and investment management- practices; all being driven by differences in the way Protestants view their religious beliefs.

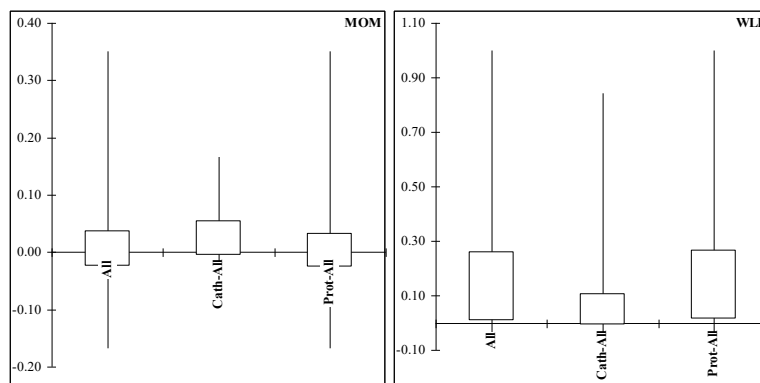
Examining the *SMB* (Small-minus-Big) factor loadings in Figure 2, we find that both Catholic and Protestant fund types are exposed to non-trivial levels of Size-related risk with the Catholic funds being slightly more sensitive to Size-based dynamics. In unreported figures that

disaggregate Catholic and Protestant funds by size, we find supporting evidence that the relatively larger amount of loading variability for Protestant funds is not due to the sample's construction (i.e. more Protestant funds lead to more Protestant fund variation) but is due to the religion-specific aspects of Protestant funds relative to Catholic funds. Thus, the absolute- and relative- representations of funds included in the sample does not impact the results themselves. It is, then, not the sample that is driving the results but an underlying process we attribute to the differences between the Catholic and Protestant religions. Detailed factor loading results available upon request.

Figure 2: Factor Regressions

The following box plots display various estimation results for OLS regressions performed on each of the 65 Christ-focused funds (12 Catholic and 53 Protestant) across the entire sample (January 2nd, 2018 to December 31st, 2024; n = 1,761), using Equation 1, where MKT proxies for systematic market risk, SMB proxies for the small minus big factor, HML proxies for the high book-to-market (i.e. value) minus low book-to-market (i.e. growth) factor, INV proxies for the conservative minus aggressive factor, MOM proxies for the momentum factor, and WLD proxies for an international investment factor. Each box plot reports the maximum, 75th percentile, 25th percentile, and minimum coefficient values.

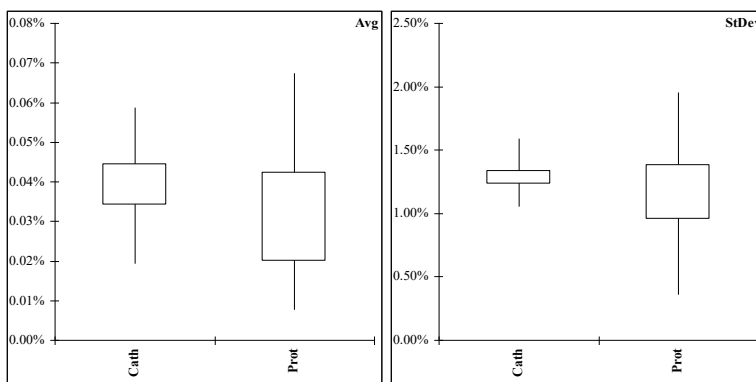


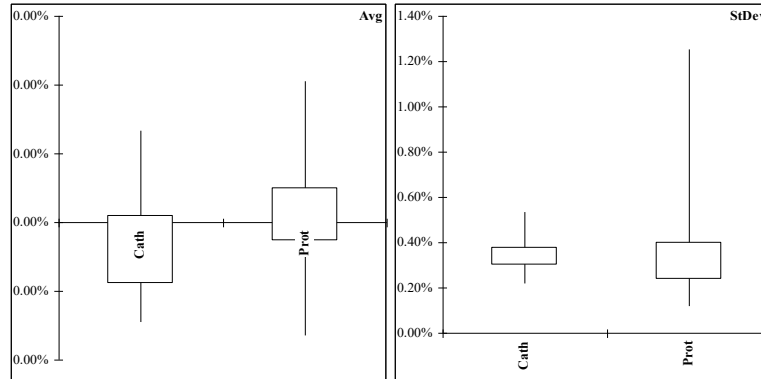


Continuing on the *HML* (i.e. High-minus-Low) factor loadings, we see that all fund groupings have a high level of coefficient variability and that the Protestant-based funds have more coefficient variability than the Catholic-based funds. The *INV* (i.e. Conservative minus Aggressive factor) and *MOM* (i.e. Momentum factor), coefficients in Figure 2 affirms the above results. Again, we find that Protestant funds tend to have a higher level of coefficient variability and that different religiously-focused funds have different patterns of factor loadings. For example, Protestant funds tend to be more sensitive to the *INV* factor indicating that, holding all else constant, Protestant funds tend to invest in companies that are more conservative in their investing styles. Further, looking at the *MOM* factor, Catholic funds tend to exhibit higher own-momentum and persistency than Protestant funds. Ending our analysis of the factor-based box plots, we examine *WLD* (i.e. World factor) loadings. Again, Figure 2 shows substantially higher levels of coefficient variation in the Protestant-centric funds relative to the Catholic-centric funds.

Figure 3: Religious Moments

The following box plots display various sample averages and standard deviations of 65 Christ-focused funds (12 Catholic and 53 Protestant) across the entire sample (January 2nd, 2018 to December 31st, 2024; $n = 1,761$). Each box plot reports the maximum, 75th percentile, 25th percentile, and minimum raw returns (top panels) and filtered returns (bottom panels) in term of these returns' average (left panels) and standard deviations (right panels), reported for different religious classifications. Filtered returns are the estimated residuals from full-sample OLS regressions using Equation 1:





Finally, we examine the impact of filtering on the various funds' returns in the Figure 3 box plots above. Here, we find that both Catholic and Protestant funds have positive daily raw returns throughout the sample, that Protestant funds have a higher degree of raw return variation, and that Catholic funds tend to have higher levels of raw returns than Protestant funds. Yet, when examining the filtered daily returns (i.e. the estimated residuals of Equation 1), we find that the raw returns results are largely reversed. That is, Catholic funds predominantly have negative risk-adjusted returns while a substantial proportion of Protestant funds still exhibit positive (risk-adjusted) returns.

Further, the 25th/75th percentile range of filtered returns is lower for Protestant funds than for Catholic funds. Again, this is a systematic difference that cannot be explained by the sample in terms of style, size, conservativeness, momentum, or location preferences; nor is this a function of unexplained risk factors not being incorporated into the model, as the Adjusted-R²s (see Figure 2) are very high for all fund groupings. Rather, Protestant funds tend to have different-asset pricing and risk-sensitivity characteristics than Catholic funds. All of this, once again, collectively continues a consistent message seen throughout all of the various factor loading box plots and associated analyses: different religiously-oriented funds exhibit different empirical outcomes that can, ultimately, only be a manifestation of their systematically-different underlying religious beliefs.

Rolling Correlation Results

As noted in the findings above in Figure 3: on average, there are systematic differences in the Catholic and Protestant fund grouping returns after filtering for a large variety of potential asset pricing drivers. In other words, there appears to be a religion-specific asset pricing factor embedded within FBI funds and that this asset pricing factor is distinct for Catholic funds relative to Protestant funds. We further explore this unique facet of FBI fund dynamics by plotting equally-weighted indexes of Catholic, Protestant, Islamic, and Israeli investment funds.

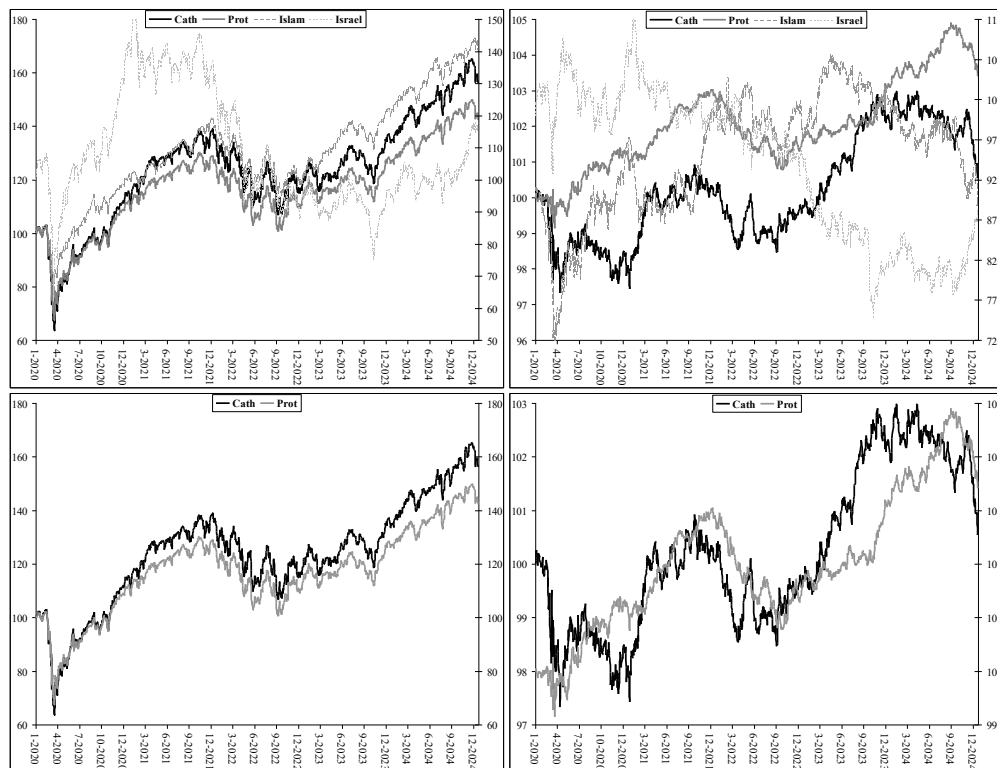
The left panels of Figure 4 (below) are for raw faith index returns and show that, with the exception of the Israeli index prior to 2021-Q1, all FBI indexes exhibit high levels of synchronicity, moving largely in lockstep fashion. Granted there are differences in the end-sample levels, yet the comovement among the funds, especially after 2021-Q1, is unmistakable. However, when examining the indexes constructed from model-filtered (Equation 1) returns in the right panels of Figure 4, the strong cross-religion index comovement breaks down. Catholic and Protestant funds are still relatively correlated but now they differ in terms of timing and magnitude. The correlations are even more different for the Islamic and Israeli model-filtered

indexes. Collectively, these results further point to religion-specific pricing dynamics that impact different funds differently. Further, adding the Islamic and Israeli indexes shows preliminary evidence that relatively-similar (largely-different) religions have relatively-similar (largely-different) religious-specific systematic risk exposure.

Honing in on this notion of religion-specific asset pricing factors that are less (more) different for less (more) different religiously-oriented funds, we plot the 90-day rolling index raw return correlations across a shortened sample in Figure 5. Figure 5 shows that Catholic/Protestant index raw return rolling correlations are the highest across the sample and exhibit the lowest level of variability. Thus, the relatively-similar Catholic and Protestant religion indexes are highly related and stable across market conditions. The Islamic/Israeli relationship, in stark contrast, is lower in magnitude across the sample and exhibits significant bouts of time variation; increasing during periods of heightened systematic market volatility. In both cases, the benefits to diversifying across religions are largely lacking: the high Catholic/Protestant correlations mean that diversification benefits are low across all market conditions whereas the crisis-heightened Islamic/Israeli correlations mean that diversification benefits are lowest when they are most needed.

Figure 4: Index Comparisons

The following figures display index values (Index = 100 on January 6th, 2020) for four, faith-based classifications: a Catholic index of 12 underlying Catholic-focused investment funds, a Protestant index of 53 Protestant-focused funds, an Islamic index of 3 Islam-focused funds, and an Israeli index of 2 Israel-focused funds across a sample shortened to accommodate the Islamic funds (January 6th, 2020 to December 31st, 2024; n = 1,256). The left figures display indexes constructed using raw returns whereas the right figures display indexes constructed from the estimated residuals of shortened-sample OLS regressions using Equation 1.



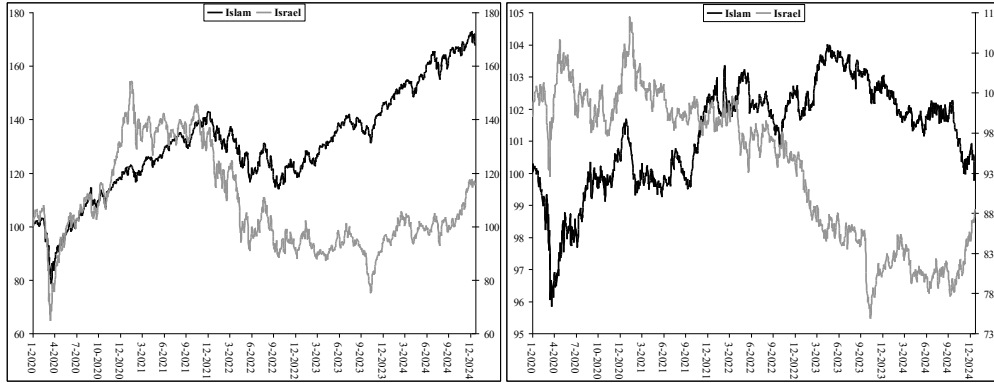
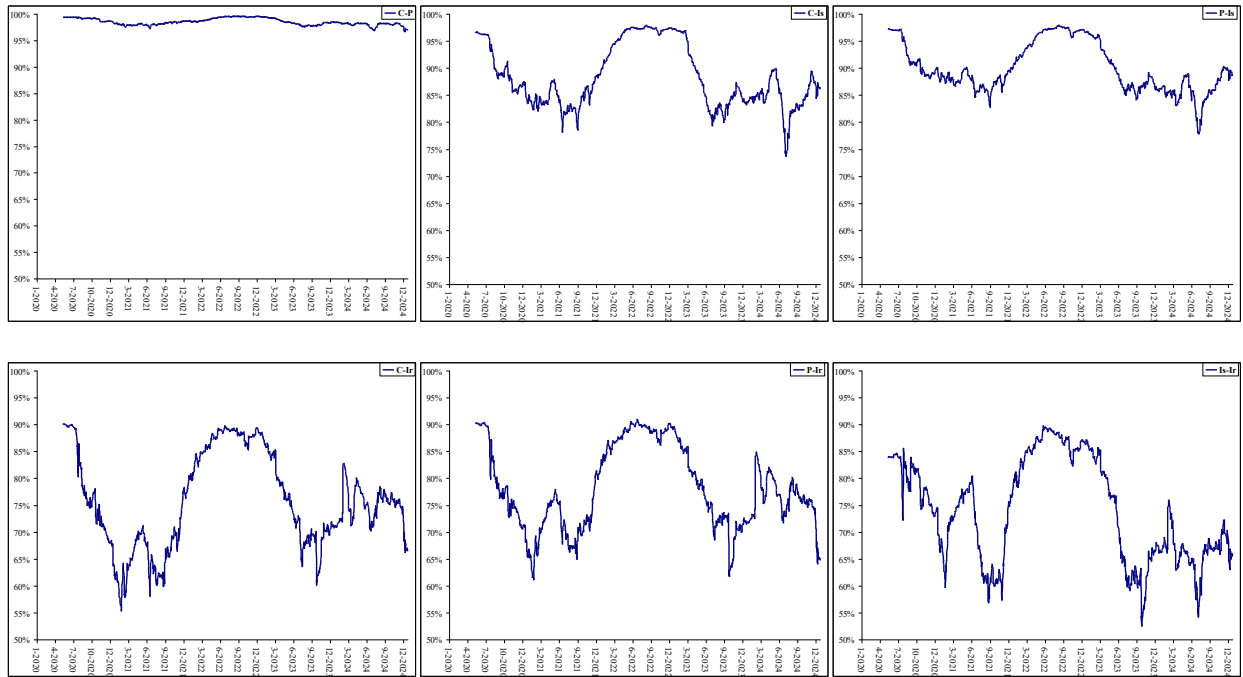


Figure 5: Raw Rolling Correlations

The following figures display the 90-day rolling correlation coefficients among the raw returns of four, faith-based indexes: a Catholic index of 12 underlying Catholic-focused investment funds (C), a Protestant index of 53 Protestant-focused funds (P), an Islamic index of 3 Islam-focused funds (Is), and an Israeli index of 2 Israel-focused funds (Ir) across a sample shortened to accommodate the Islamic funds (January 6th, 2020 to December 31st, 2024; $n = 1,256$).



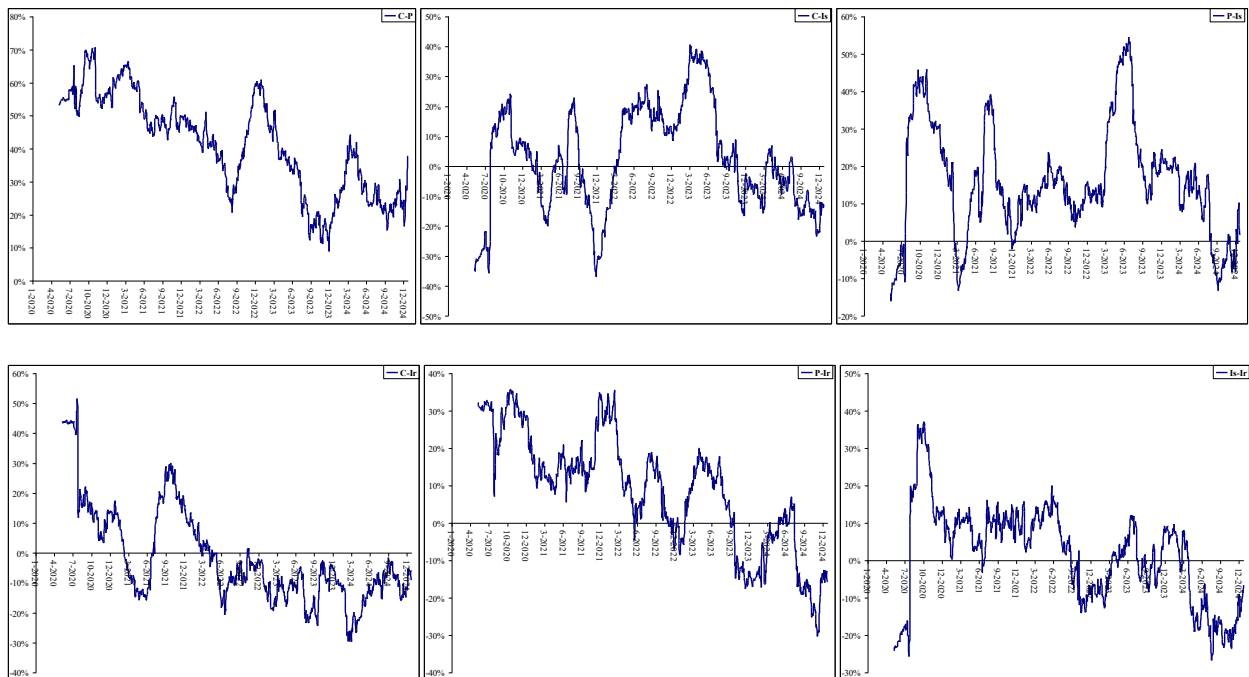
Evidence of religious-specific asset pricing factors becomes even more pronounced when looking at the model-filtered (Equation 1) index return correlations in Figure 6. Similar to what we found in Figure 1, the time-varying filtered index return correlations show that accounting for common style, size, conservatism, momentum, and location factors significantly reduces cross-religion correlations. This indicates that perceptions of commonality among faith funds is largely an aberration driven by common asset pricing factors. In addition to what is found in Figure 1,

Figure 6 highlights that the time varying nature of cross-religion correlations is significant and trends in two separate ways. First, cross-religion correlations tend to increase during periods of relative market stress. This crisis effect is found even for the Catholic/Protestant pair that, in the raw correlations, was universally high and time invariant across the sample.

Second, and with the exception of the Catholic/Islamic and Protestant/Islamic correlations for the first half of the sample, the time-varying correlations are negatively trending across time, indicating that cross-religious distinctions among faith funds is increasing. Inferentially, this could mean that faith fund managers are more adept at filtering investments according-to and in compliance-with their clients' respective religious beliefs. In accordance with the filtered religious indexes in Figure 4, and the raw religious correlations in Figure 6, we again find that relatively more (less) similar religious indexes are more (less) correlated. This last result suggests that different religious-focused funds will have systematically-different screening behaviors and, hence, different systematic religious pricing impacts.

Figure 6: Filtered Rolling Correlations

The following figures display the 90-day rolling correlation coefficients among the filtered returns of four, faith-based indexes: a Catholic index of 12 underlying Catholic-focused investment funds (C), a Protestant index of 53 Protestant-focused funds (P), an Islamic index of 3 Islam-focused funds (Is), and an Israeli index of 2 Israel-focused funds (Ir) across a sample shortened to accommodate the Islamic funds (January 6th, 2020 to December 31st, 2024; n = 1,256). All returns used to calculate the various rolling correlations are the estimated residuals from shortened sample OLS regressions using Equation 1.



Interfaith Index Regressions

We end our empirical analysis by elaborating into this last finding: that different funds have different, religion-specific asset pricing factors. We do this by modeling equally-weighted

Catholic fund index returns with the factor filtration model of Equation 1 and equally-weighted Protestant fund index returns (Equation 2).

Table 2: Catholic vs. Protestant Religious Regressions

The following table reports the estimation results of two OLS regressions performed on a Catholic Index of 12 Catholic-focused investment funds across a shortened sample (January 6th, 2020 to December 31st, 2024; $n = 1,256$). The first estimation includes an index of 53 Protestant-focused investment funds (Equation 2) whereas the second estimation only includes six risk factors (Equation 1). For both estimations, *MKT* proxies for systematic market risk, *SMB* proxies for the small minus big factor, *HML* proxies for the high book-to-market (i.e. value) minus low book-to-market (i.e. growth) factor, *INV* proxies for the conservative minus aggressive factor, *MOM* proxies for the momentum factor, and *WLD* proxies for an international investment factor.

	Coef	Se	t-Stat	p-Value	Coef	Se	t-Stat	p-Value
Alpha	0.000	0.000	-0.473	0.637	0.000	0.000	-1.355	0.176
MKT	0.486	0.026	18.772	0.000	0.868	0.011	80.563	0.000
SMB	0.164	0.009	18.600	0.000	0.263	0.007	38.690	0.000
HML	0.052	0.005	10.848	0.000	0.064	0.005	12.436	0.000
INV	-0.076	0.018	-4.143	0.000	-0.002	0.020	-0.122	0.903
MOM	0.001	0.006	0.185	0.853	0.003	0.007	0.395	0.693
WLD	-0.010	0.011	-0.907	0.364	0.120	0.008	14.363	0.000
Prot	0.550	0.034	15.937	0.000	--	--	--	--
	<i>Adjusted R² = 0.988</i>				<i>Adjusted R² = 0.986</i>			

From Table 2, we find that Catholic index returns are well-explained by the *MKT*, *SMB*, *HML*, and *WLD* factors. Each of these loadings are statistically significant at less than the 1% level and, with the exception of the *HML* loading, are economically significant. The sum of significant estimated coefficients is 1.315 and the Adjusted- R^2 is a lofty 0.986. When augmenting the base factor model with equally-weighted Protestant index returns (i.e. *Prot_t*), we find in Table 2 that the Adjusted- R^2 is slightly increased to almost unitary, all previously statistically significant coefficients are of similar statistical significance and of the same coefficient sign (with the exception of *WLD*), and the *INV* factor has become significant. However, the economic significance of the statistically significant factor loadings has decreased by about a third on average. This drop in the factors' economic significance is explained by the Protestant factor being significant both statistically and economically. In fact, the *Prot* coefficient of 0.550 is stronger than any other pricing factor and is almost equivalent to the sum of other, statistically significant factor loadings (0.625).

Combined, Table 2's results clearly indicate that, first, Protestant funds have a common pricing component with Catholic funds that transcends traditional asset pricing factors. Second, the Protestant pricing factor does not dominate other pricing factors nor is the Protestant factor of unitary value. Thus, third, Catholic and Protestant funds share a common, but not identical, asset pricing feature; there is still religion-specific variation that exists with the Catholic index returns that cannot be accounted for by the Protestant index returns. Put differently: despite their supposed religious similarities, Catholic and Protestant funds are treated as being different by the market.

Tables 3 and 4 largely back up the results seen above when modeling Catholic and Protestant equally-weighted index returns as a function of the Equation 1 asset pricing factors as well as equally-weighted Islamic and Israeli index returns.

Table 3: Catholic Index Religious Regressions

The following table reports the estimation results of two OLS regressions performed on a Catholic Index of 12 Catholic-focused investment funds across a sample shortened to accommodate two non-Christian indexes (January 6th, 2020 to December 31st, 2024; $n = 1,256$). The first estimation includes the Islam and Israel equity indexes (Equation 3) whereas the second estimation only includes six risk factors (Equation 1). For both estimations, MKT proxies for systematic market risk, SMB proxies for the small minus big factor, HML proxies for the high book-to-market (i.e. value) minus low book-to-market (i.e. growth) factor, INV proxies for the conservative minus aggressive factor, MOM proxies for the momentum factor, and WLD proxies for an international investment factor.

	Coef	Se	t-Stat	p-Value	Coef	Se	t-Stat	p-Value
Alpha	0.000	0.000	-1.276	0.202	0.000	0.000	-1.355	0.176
MKT	0.882	0.018	48.909	0.000	0.868	0.011	80.563	0.000
SMB	0.253	0.008	32.134	0.000	0.263	0.007	38.690	0.000
HML	0.068	0.007	10.178	0.000	0.064	0.005	12.436	0.000
INV	0.006	0.020	0.282	0.778	-0.002	0.020	-0.122	0.903
MOM	0.001	0.007	0.212	0.832	0.003	0.007	0.395	0.693
WLD	0.117	0.009	13.656	0.000	0.120	0.008	14.363	0.000
Islam	-0.030	0.022	-1.364	0.173	--	--	--	--
Israel	0.015	0.006	2.624	0.009	--	--	--	--
	<i>Adjusted R² = 0.986</i>				<i>Adjusted R² = 0.986</i>			

Here, we again see that the *MKT*, *SMB*, *HML*, *WLD*, and *INV* (Protestant-only) factor loadings are both statistically- and economically- significant. Further, the sum of estimated coefficients is large for both Equation 1 estimations (Catholic = 1.315 and Protestant = 1.269) and the Adjusted-R²s are very high. Furthermore, and for both Equation 3 estimations, the Islamic factor loading is statistically insignificant while the Israeli factor loading is statistically significant at the less than 1% level. However, and unlike the Equation 2 estimations found in Table 2, the Equation 3 estimations found in Tables 3 and 4 indicate that adding religiously-different index returns does not substantially alter or improve the Catholic or Protestant models; the Adjusted-R²s and coefficient sums are largely unchanged when adding the cross-religious returns.

Further, the Israeli factor loading for both the Catholic and Protestant estimations is lower in economic significance (individually) than any other statistically significant coefficient by an average of 1-to-21.4 and 1-to-8.68, respectively. Collectively, the Israeli coefficient is 1.17% and 2.30% the size of the other significant factor loadings added together, respectively. Thus, while there is a small, cross-religious impact of Israeli funds onto both Catholic and Protestant funds, the impact is extremely low and not of practical economic significance. This is made all the more clear in that the Protestant coefficient from Table 2 is approximately 35.68 times larger than the Israeli coefficient in Table 3. For the Islamic factor loadings, we find that they are not statistically different from zero in either of the Catholic or Protestant estimations. While caution must be applied in comparing these insignificant coefficient estimates with statistically significant coefficients, we note that the Islamic coefficients are smaller in absolute magnitude than any of the significant coefficients (individually or combined) and are smaller in relative magnitude than the Israeli index coefficient. These results indicate that different religions have different, cross-religious asset pricing effects but that more (less) similar religions have a higher (lower) cross-religious impact. Put differently: different religious FBI funds have different

religion-specific asset pricing factors, but that more (less) similar religions will have more (less) similar religion-specific asset pricing factors.

Table 4: Protestant Index Religious Regressions

The following table reports the estimation results of two OLS regressions performed on a Protestant Index of 53 Protestant-focused investment funds across a sample shortened to accommodate two non-Christian funds (January 6th, 2020 to December 31st, 2024; $n = 1,256$). The first estimation includes the Islam and Israel equity indexes (Equation 3) whereas the second estimation only includes six risk factors (Equation 1) For both estimations, MKT proxies for systematic market risk, SMB proxies for the small minus big factor, HML proxies for the high book-to-market (i.e. value) minus low book-to-market (i.e. growth) factor, INV proxies for the conservative minus aggressive factor, MOM proxies for the momentum factor, and WLD proxies for an international investment factor.

	Coef	Se	t-Stat	p-Value	Coef	Se	t-Stat	p-Value
Alpha	0.000	0.000	-2.123	0.034	0.000	0.000	-2.244	0.025
MKT	0.673	0.013	50.607	0.000	0.695	0.008	86.187	0.000
SMB	0.161	0.006	27.840	0.000	0.181	0.005	35.568	0.000
HML	0.040	0.005	8.081	0.000	0.023	0.004	5.858	0.000
INV	0.140	0.015	9.591	0.000	0.134	0.015	9.195	0.000
MOM	0.001	0.005	0.124	0.901	0.003	0.005	0.550	0.582
WLD	0.227	0.006	35.999	0.000	0.236	0.006	37.876	0.000
Islam	0.018	0.016	1.084	0.279	--	--	--	--
Israel	0.029	0.004	6.606	0.000	--	--	--	--
	<i>Adjusted R² = 0.990</i>				<i>Adjusted R² = 0.989</i>			

V. CONCLUDING DISCUSSIONS

This paper has provided a detailed review of the existing literature on Socially Responsible Investing (SRI) and Faith Based Investing (FBI), a systematic analysis of various faith traditions' core beliefs, and a comprehensive empirical analysis of individual- and index- level FBI funds. In and throughout all of this, we have shown unmistakable and consistent evidence that differences exist among the three primary FBI faith-classifications: Roman Catholicism, Biblical Christianity as proxied by Protestantism, and Islam. With each of the three modes of analysis we use, the extant literature, the underlying beliefs, and the empirical examination, a clear picture emerges: different faiths have different underlying beliefs, that these beliefs are reflected in FBI funds' screening and investment behavior, and that all of this is reflected in empirical market outcomes.

Based on what is presented in this study, we make a call for academics, investment managers, the investors themselves, and financial regulators to be aware of and take consistent action regarding these faith traditions' differences. We encourage all to resist the siren call of easy ecumenicism which attempts to blend faiths and, especially in an investments context, can lead to suboptimal results for all. For academic researchers: erroneous ecumenicism glosses over potentially rich results that could be found when parsing out effects by faith. While certainly done without malice or intent, academics' past willingness to conflate all faiths as one, or treat one faith as all, subtly paints a picture that is not accurate either from a faith-perspective or from the perspective of financial market outcomes.

Amazing insights have been gleaned throughout the years from studies that have marginalized, however implicitly and passively, the differences in the faiths. Now is the

opportunity for academics to revisit the existing literature for opportunities to parse the results out according to various faiths. Further, and as has been noted throughout this analysis, there are clear avenues for furthering the literature beyond simply distinguishing among the faiths. Opportunities abound, including: examining intradenominational distinctions; time-varying behavior among secular funds, secular-SRIs, and FBIs; exploring the notion of systematic faith-based asset pricing factors; and more.

For investment management professionals: this study provides clear evidence that specializing one's management practice or funds toward a particular faith, is a wise strategy. In doing so, faith-focused clients have their faith-idiosyncratic demand addressed through negative screening approaches and, with mixed results, active advocacy and activism. As is seen in the literature, investors have faith-specific components as part of their overall, multi-dimensional investment utility functions and show a strong willingness to pay for tailor-managed funds. Further, as seen in the rolling, time-varying conditional correlations, both Catholic- and Protestant- focused FBIs are increasingly specializing within (and apart from) Christ-centric (non-Christ-centric) FBIs. In short, this study provides a good rationale for investment professionals to avoid erroneous ecumenicism and that, with constant vigilance, they can continue to provide unique investment opportunities to their clients.

Yet, investment management professionals should also be cautious of efforts, driven indirectly by ecclesiastical efforts and supranational organizations, and more directly by these efforts' industry-facing advocacy groups, to soften faith distinctions. Investment managers are called to maintain their fiduciary duty to their clients' financial and non-financial utility maximizing needs; not allowing interests and encouragements, from unrelated faiths or even human secular concerns, to get in the way of ensuring that faith-specific clients are being served with faith-specific efforts. Further, given the understandably wide net these ecumenical efforts are casting, it is not surprising that an equally wide ranging set of participants are being brought together. Yet, and not exclusively for US-based, Biblically-literal Christians, associations with potentially spiritually misaligned groups warrant particular caution, whether these associations are with the human-secularist United Nations or Open Society Foundation; the increasingly-globalist Roman Catholic Church and its ecumenical endeavors; or even Christ marginalizing, optional, or denying faith traditions such as Universal Unitarianism and others.

For faith-driven investors: they should, also, be aware of and respond appropriately to investment managers' faith-specific offerings; ensuring that the managers they hire, or funds they participate in, are unambiguously aligned with their particular faith. While this study has found that similarities do exist, for example, among Catholic- and Protestant- focused funds, it is also clear that these similarities are not one-to-one, the performance of these funds differ, and these funds are driven by systematically different investment management behaviors and philosophies. Catholic-focused funds and investment management professionals may or may not be suitable for Protestant-focused investors, and vice versa; investment decisions should be made on a case-by-case basis and only after considerable research and discerning reflection. Faith-oriented investors are encouraged to have conversations with their investment professionals and fund providers about some of the issues addressed in this paper and balance financial, risk and return considerations with non-financial faith, activism, and ethical concerns.

Finally, for financial regulators: we encourage consistent vigilance in ensuring that all market participants are clearly informed of the distinctions within and among faith-focused funds, what is being offered to investors, and what is not. Based on the results on this study, and only looking at fund managers' abilities to distinctly manage distinctly different faith needs, it

appears that many Catholic- and Protestant- oriented funds are doing exactly that: meeting the faith-specific needs of their clients. Still, we encourage regulators to ensure that FBI managers and funds are fully disclosing their spiritual alignment, related screening methodologies, and any activist positions taken. Regulators are also encouraged to ensure that investors are well aware that FBI fund financial performance may, depending on market conditions and the fund at hand, have different raw- and risk adjusted- return characteristics than market- and secular-benchmarks; but also that differing financial performance is not necessarily indicative of a breach of fiduciary duty of those managers and funds charged with co-maximizing faith-based, non-financial utility with traditional financial utility.

We pray that this study serves as an enlightening, edifying waypoint on the path of future studies.

References

- Adams, J., & Ahmed, P. 2013. The performance of faith-based funds. *Journal of Investing*, 22(4), 83-92.
- Adler, T., & Kritzman, M. 2008. The cost of socially responsible investing. *Journal of Portfolio Management*, 35(1), 52-56.
- Areal, N., Cortez, M., & Silva, F. 2010. Investing in mutual funds: Does it pay to be a sinner or a saint in times of crisis? University of Minho School of Economics and Management Working Paper.
- Barnett, M., & Salomon, R., 2006. Beyond dichotomy: The curvilinear relationship between social responsibility and financial performance. *Strategic Management Journal*, 27(11), 1101-1122.
- Bartke, S., & Schwarze, R. 2008. Risk-averse by nation or by religion? Some insights on the determinants of individual risk attitudes. German Socio-Economic Panel (SOEP) Working Paper.
- Bauer, R., Derwall, R., & Otten, R. 2007. The ethical mutual fund performance debate: New evidence from Canada. *Journal of Business Ethics*, 70(2), 111-124.
- Bauer, R., Otten, R., & Rad, A. 2006. Ethical investing in Australia: Is there a financial penalty? *Pacific-Basin Finance Journal*, 14(1), 33-48.
- Beer, F., Estes, J., & Deshayes, C. 2014. Performance of the faith and ethical investment products: A comparison before and after the 2008 meltdown. *Financial Services Review*, 23(2), 151-167.
- Blau, B. 2017. Religiosity and the volatility of stock prices: A cross-country analysis. *Journal of Business Ethics*, 144(3), 609-621.
- Boasson, E., Boasson, V., & Cheng, J. 2006. Investment principles and strategies of faith-based funds. *Managerial Finance*, 32(10), 837-845.
- Boasson, V., Cheng, C., & Boasson, E. 2004. Are investment managers investing ethically at a disadvantage? *Journal of Applied Management and Entrepreneurship*, 9(4), 56-65.
- Bollen, N. 2007. Mutual fund attributes and investor behavior. *Journal of Financial and Quantitative Analysis*, 42(3), 683-708.
- Boone, J., Khurana, I., & Raman, K. 2013. Religiosity and tax avoidance. *Journal of the American Taxation Association*, 35(1), 53-84.
- Borgers, A., & Pownall, R. 2014. Attitudes towards socially and environmentally responsible investment. *Journal of Behavioral and Experimental Finance*, 1, 27-44.

- Callen, J., & Fang, X. 2015. Religion and stock price crash risk. *Journal of Financial and Quantitative Analysis*, 50(1), 169-195.
- Caner, E. 2015. Islam. *World Religions and Cults: Volume 1*. New Leaf Publishing Group, Green Forest Arkansas.
- Carhart, M. 1997. On persistence in mutual fund performance. *Journal of Finance*, 52(1), 57-82.
- Carosella, N., Rodriguez, J., Williams, S., Nawrocki, D., & Doh, J. 2012. The return-to-risk performance of socially responsible investing according to Catholic values. *Journal of Investing*, 21(4), 47-58.
- Chang, E., & Witte, D. 2010. Performance evaluation of U.S. socially responsible mutual funds: Revisiting doing good and doing well. *American Journal of Business*, 25(1), 9-21.
- Chen, J., & Diaz, J. 2012. Spillover and leverage effects of faith-based exchange-traded funds. *Journal of Business Policy and Research*, 7(2), 1-12.
- Chen, H., Huang, H., Lobo, G., & Chong, W. 2016. Religiosity and the cost of debt. *Journal of Banking and Finance*, 70(9), 70-85.
- Chircop, J., Johan, S., & Tarsalewska, M. 2020. Does religiosity influence venture capital investment decisions? *Journal of Corporate Finance*, 62, 1-14.
- Clementino, E., & Perkins, R. 2020. How do companies respond to environmental, social, and governance (ESG) ratings? Evidence from Italy. *Journal of Business Ethics*, 171(2), 379-397.
- Cortez, M., Silva, F., & Areal, N. 2009. The performance of European socially responsible funds. *Journal of Business Ethics*, 87(4), 573-588.
- Czerwonka, M. 2014. The influence of religion on socially responsible investing. *Journal of Religion and Business Ethics*, 3(1), 1-10.
- Davis, M., Ghouli, W., & Thies, C. 2017. The performance of faith-based mutual funds: A matched-pair approach. *Journal of Applied Business and Economics*, 19(7), 75-83.
- Derwall, J., Koedijk, K., & Horst, J. 2011. A tale of values-driven and profit-seeking social investors. *Journal of Banking and Finance*, 35(8), 2137-2147.
- Diaz, J. 2000. Religion and gambling in Sin-City: A statistical analysis of the relationship between religion and gambling patterns in Las Vegas residents. *Social Science Journal*, 37(3), 453-458.
- diBartolomeo, D., & Kurtz, L. 1999. Managing risk exposures of socially screened accounts. Northfield Working Paper.
- Diener, J. 2022. Impact case or impact washing? An analysis of investors' strategies to influence corporate behavior. *Sustainability Accounting, Management, and Policy Journal*, 14(5), 1002-1021.
- Diener, J., & Habisch, A. 2022a. Developing an impact-focused typology of socially responsible fund providers. *Journal of Risk and Financial Management*, 15(7), 298-316.
- Diener, J., & Habisch, A. 2022b. God's stewards: A global overview of Christian-influenced mutual fund providers. *Journal of Risk and Financial Management*, 15(12), 547-571.
- Diltz, D. 1995. Does social screening affect portfolio performance? *Journal of Investing*, 1, 64-69.
- Dion, M. 2009. Christian mutual funds, codes of ethics, and corporate illegalities. *International Journal of Social Economics*, 36(9), 916-929.
- Dunfee, T. 2003. Social investing: Mainstream or backwater? *Journal of Business Ethics*, 43(3), 247-252.

- Ellison, C., & Nybrotten, K. 1999. Conservative Protestantism and opposition to state-sponsored lotteries: Evidence from the 1997 Texas poll. *Social Science Quarterly*, 80(2), 356-369.
- Enete, S., & Kiss, E. 2018. How negative screening according to Christian principles influence stock returns. 2018 Academic Research Colloquium for Financial Planning and Related Disciplines.
- Fama, E., & French, K. 2015. A five-factor asset pricing model. *Journal of Financial Economics*, 116(1), 1-22.
- Fairbanks, D. 1977. Religious forces and "morality" policies in the American states. *Western Political Quarterly*, 30(3), 411-417.
- Ferruz, L., Muñoz, F., & Vargas, M. 2012. Managerial abilities: Evidence from religious mutual fund managers. *Journal of Business Ethics*, 105(4), 503-517.
- Gao, L., Wang, Y., & Zhao, J. 2017. Does local religiosity Affect organizational risk-taking? Evidence from the hedge fund industry. *Journal of Corporate Finance*, 47, 1-22.
- Geczy, C., Stambaugh, R., & Levin, D. 2005. Investing in socially responsible mutual funds. Working Paper.
- Gellman, L. 2013. Investing as a religious practice: Faith-based mutual funds incorporate ethical values in selecting securities to own. WSJ.com, November 13th, 2013.
- Gharbi, I., Hamed-Sidhom, M., Hussainey, K., & Ganouati, J. 2020. Religiosity and financial distress in U.S. firms. *International Journal of Finance and Economics*, 26(3), 3902-3915.
- Ghoul, W., & Karam, P. 2007. MRI and SRI mutual funds: A comparison of Christian, Islamic (morally responsible investing), and socially responsible investing (SRI) mutual funds. *Journal of Investing*, 16(2), 96-102.
- Gutsche, G. 2019. Individual and regional Christian religion and the consideration of sustainable criteria in consumption and investment decisions: An exploratory econometric analysis. *Journal of Business Ethics*, 157(4), 1155-1182.
- Halek, M., & Eisenhauer, J. 2001. Demography of risk aversion. *Journal of Risk and Insurance*, 68(1), 1-24.
- Harjoto, M., Rossi, F. 2019. Religiosity, female directors, and corporate social responsibility for Italian listed companies. *Journal of Business Research*, 95, 338-346.
- Hilary, G., & Hui, H.W. 2009. Does religion matter in corporate decision making in America? *Journal of Financial Economics*, 93(3), 455-473.
- Hillman, A., & Potrafke, N. 2018. Economic freedom and religion: An empirical investigation. *Public Finance Review*, 46(2), 249-275.
- Hoffman, J. 2000. Religion and problem gambling in the U.S. *Review of Religious Research*, 41(4), 488-509.
- Hood, M., Nofsinger, J., & Varma, A. 2013. Conservation, discrimination, and salvation: Investors' social concerns in the stock market. *Journal of Financial Services Research*, 45(1), 5-37.
- Inskeep, K. 1992. Views on social responsibility: The investment of pension funds in the evangelical Lutheran Church in America. *Review of Religious Research*, 33(3), 270-282.
- Ito, Y., Managi, S., & Matsuda, A. 2013. Performances of socially responsible investment and environmentally friendly funds. *Journal of the Operational Research Society*, 64(11), 1583-1594.
- Johnson, C., & Meier, K. 1990. The wages of sin: Taxing America's legal vices. *Western Political Quarterly*, 43(3), 577-595.

- Junkus, J., & Berry, T. 2015. Socially responsible investing: A review of the critical issues. *Managerial Finance*, 41(11), 1176-1201.
- Kim, I., & Venkatachalam, M. 2006. Are sin stocks paying the price for their accounting sins? *Journal of Accounting, Auditing, and Finance*, 26(2), 415-442.
- Kinder, P., & Domini, A. 1997. Social screening: Paradigms old and new. *Journal of Investing*, 6(4), 12-19.
- Kreander, N., McPhail, K., & Molyneaux, D. 2004. God's fund managers: A critical study of stock market investment practices of the Church of England and UK Methodists. *Accounting, Auditing, and Accountability Journal*, 17(3), 408-441.
- Kumar, A. 2009. Who gambles in the stock market? *Journal of Finance*, 64(6), 1889-1933.
- Kumar, A., & Page, J. 2013. Deviations from norms and informed trading. *Journal of Financial and Quantitative Analysis*, 49(4), 1005-1037.
- Kumar, A., Page, J., & Spalt, O. 2011. Religious beliefs, gambling attitudes, and financial market. *Journal of Financial Economics*, 102(3), 671-708.
- Kurtz, L., & diBartolomeo, D. 2005. The KLD Catholic values 400 index. *Journal of Investing*, 14(3), 101-104.
- Lee, D., Humphrey, J., Benson, K., & Ahn, J. 2010. Socially responsible investment fund performance: The impact of screening intensity. *Accounting and Finance*, 50(2), 351-370.
- Leon, A., & Pfeifer, C. 2017. Religious activity, risk-taking preferences, and financial behaviour: Empirical evidence from German survey data. *Journal of Behavioral and Experimental Economics*, 69(1), 99-107.
- Liston, D., & Soydemir, G. 2010. Faith-based and sin portfolios: An empirical inquiry into norm-neglect vs. norm-conforming investor behavior. *Managerial Finance*, 36(10), 876-885.
- Louche, C., Arenas, D., & van Cranenburgh, K. 2012. From preaching to investing: Attitudes of religious organizations towards responsible investment. *Journal of Business Ethics*, 110(3), 310-320.
- Lyn, E., & Zychowicz, E. 2010. The impact of faith-based screens on investment performance. *Journal of Investing*, 19(3), 136-143.
- Mao, L., Lu, C., Sun, G., Zhang, C., & Guo, C. 2024. Regional culture and corporate finance: A literature review. *Humanities and Social Sciences Communications*, 11(59), 1-12.
- Markowitz, H. 1952. Portfolio selection. *Journal of Finance*, 7(1), 77-91.
- McGuire, S., Omer, T., & Sharp, N. 2012. The impact of religion on financial reporting irregularities. *Accounting Review*, 87(2), 645-673.
- Meunier, L., & Ohadi, S. 2023. Exclusion strategy in socially responsible investment: One size does not fit all. *Journal of Behavioral and Experimental Finance*, 39, 1-19.
- Miller, A., & Hoffmann, J. 1995. Risk and religion: An explanation of gender differences in religiosity. *Journal for the Scientific Study of Religion*, 34(1), 63-75.
- Montenegro, T. 2017. Religiosity and corporate financial reporting: Evidence from a European country. *Journal of Management, Spirituality, and Religion*, 14(1), 48-80.
- Mortenson, T. 2015. How is Roman Catholicism different? *World Religions and Cults: Volume I*. New Leaf Publishing Group, Green Forest Arkansas.
- Naber, M. 2001. Catholic investing: The effects of screens on financial returns. *Journal of Investing*, 10(4), 51-71.
- Niszczota, P., & Bialek, M. 2021a. The effect of gender and parenting daughters on judgments of morally controversial companies. *Plos One*, 16(12), 1-17.

- Niszczoła, P., & Bialek, M. 2021b. Women oppose sin stocks more than men do. *Finance Research Letters*, 41, 1-6.
- Nofsinger, J., & Varma, A. 2014. Socially responsible funds and market crises. *Journal of Banking and Finance*, 48(9), 180-193.
- Noussair, C., Trautmann, S., van de Kuilen, G., & Vellekoop, N. 2013. Risk aversion and religion. *Journal of Risk and Uncertainty*, 47(2), 165-183.
- Page, J., & Spalt, O. 2011. Religious beliefs, gambling attitudes, and financial market outcomes. *Journal of Financial Economics*, 102(3), 671-708.
- Reeves, M. 2010. *The unquenchable flame: Discovering the heart of the Reformation*. B&H Publishing, Nashville Tennessee.
- Renneboog, L., Horst, J.T., & Zhang, C. 2008a. The price of ethics and stakeholder governance: The performance of socially responsible mutual funds. *Journal of Corporate Finance*, 14(3), 302-322.
- Renneboog, L., Horst, J.T., & Zhang, C. 2008b. Socially responsible investments: Institutional aspects, performance, and investor behaviour. *Journal of Banking and Finance*, 32(9), 1723-1742.
- Renneboog, L., Horst, J.T., & Zhang, C. 2011. Is ethical money financially smart? Nonfinancial attributes and money flows of socially responsible investment funds. *Journal of Financial Intermediation*, 20(4), 562-588.
- Renneboog, L., & Spaenjers, C. 2012. Religion, economic attitudes, and household finance. *Oxford Economic Press*, 64(1), 103-127.
- Rossi, F., Barth, J., & Sun, Y. 2019. Does religiosity affect the dividend payouts of Italian-listed companies? *Journal of Management, Spirituality, and Religion*, 16(4), 348-371.
- Rudd, A. 1981. Social responsibility and portfolio performance. *California Management Review*, 23(4), 55-61.
- Russell, H., Ariail, D., Smith, K., & Smith, L. 2023. Religiosity and risk: Association of Judeo-Christian ethicality with a sustainable business environment. *Journal of Risk and Financial Management*, 16(9), 394-413.
- Salaber, J. 2007. The determinants of sin stock returns: Evidence on the European market. Paris December 2007 Finance International Meeting AFFI-EUROFIDAI Paper.
- Salaber, J. 2009. Sin stock returns over the business cycle. University of Bath School of Management Working Paper.
- Salaber, J. 2013. Religion and returns in Europe. *European Journal of Political Economy*, 32, 149-160.
- Sauer, David (1997). The impact of social-responsibility screens on investment performance: Evidence from the Domini 400 Social Index and Domini Equity Mutual Fund. *Review of Financial Economics*, 6(2), 137-149.
- Schwartz, M. 2003. The "ethics" of ethical investing. *Journal of Business Ethics*, 43(3), 195-213.
- Shu, T., Sulaeman, J., & Yeung, E. 2012. Local religious beliefs and mutual fund risk-taking behaviors. *Management Science*, 58(10), 1779-1796.
- Slager, R., & Chapple, W. 2015. Carrot and stick? The role of financial market intermediaries in corporate social performance. *Business and Society*, 55(3), 398-426.
- Smith, J., & Smith, L.R. 2016. Socially responsible investing by universities and colleges. *Financial Management*, 45(4), 877-922.
- Statman, M. 2004. What do investors want? *Journal of Portfolio Management*, 30(5), 153-161.
- Statman, M. 2005. The religions of social responsibility. *Journal of Investing*, 14(3), 16-22.

- Statman, M., & Glushkov, D. 2009. The wages of social responsibility. *Financial Analysts Journal*, 65(4), 33-46.
- Stulz, R., & Williamson, R. 2003. Culture, openness, and finance. *Journal of Financial Economics*, 70(3), 313-349.
- Teoh, S., Welch, I., & Wazzan, P. 1999. The effect of socially activist investment policies on the financial markets: Evidence from the South African boycott. *Journal of Business*, 72(1), 35-89.
- The King James Study Bible, Second Edition*. 2013. Edward Hindson eds. Thomas Nelson, Nashville Tennessee.
- Vrecko, D., & Langer, T. 2013. What are investors willing to pay to customize their investment product? *Management Science*, 59(8), 1855-1870.
- Wan-Ni, L. 2012. Faith matters? A closer look at the performance of belief-based equity investments. *Journal of Asset Management*, 13(6), 421-436.

Appendix A: Scriptural Foundations

This study would not be complete without a detailed examination of Biblically-literal Christianity, Roman Catholicism, and the Islamic tradition. While the purpose of this paper is not to apologetically hammer out the nuanced differences of various religions, our cursory treatment should be more than sufficient to show that Biblically-literal Christianity, Roman Catholicism, and Islam are different and, consequently, should be treated differently by investment market participants. All Bible references within this Appendix are from the King James Version of the Bible where references are formatted as "Book Chapter:Verse".

Using the perspective of Biblically-literal Christianity as the backdrop for this discussion, mankind was cursed by God for the rebellious act of the first man, Adam, in the Garden of Eden (Genesis 3). From there, this physical- and spiritual- curse extended to all mankind (Romans 5:12), starts at birth for all (Psalm 51:5), and extends throughout one's life; even the Apostle Paul battled his sin nature during his life (e.g. Romans 7:19) and Mary praised her Son as her savior (Luke 1:46-47). Each person's sin nature essentially extends the original Adamic rebellion against God, manifests in different ways throughout their lives (e.g. Galatians 5:19-21; Romans 1:29-31; 1st Corinthians 6:9-10; 2nd Timothy 3:1-5), and separates a person from God (Isaiah 59:2). This freewill sin nature (Proverbs 8:36; James 1:13) extends to all humans (Romans 3:10-11,18,23), covers even the most minor of transgressions against God and His commandments (James 2:10), and includes all people, regardless of their knowledge, or lack thereof, of Christianity, Jesus, salvation, and other key Biblical concepts (Romans 1:19-20; 2:15-16; 3:20).

In Biblical Christianity, the consequence of being in sin at the point of death is temporary confinement in torment (2nd Peter 2:9; Luke 16:22-23), eventual bodily resurrection (Acts 24:15), judgment based on one's deeds throughout their life (Revelation 20:11-12), being found lacking due to the righteous inadequacy of one's deeds (Isaiah 64:6), and being cast into the lake of fire for an eternity of torment (Revelation 20:14-15; 14:10; 2nd Thessalonians 1:8-9). Concepts of reincarnation and "do overs" are inferentially denied in Biblical Christianity (Hebrews 9:27) as are concepts of soul sleep (Luke 16:22-23) and annihilationism (Revelation 20:10); death brings about immediate and eternally lasting results.

Yet, despite the assurance of one having a sin nature throughout one's life and the dire, eternal consequences thereof, Biblical Christianity also teaches that God has provided a way out: payment for one's sin by the atoning work of Jesus Christ. Specifically, while the Biblically-literal God is of unitary *essence* (Deuteronomy 6:4; 1st Corinthians 8:6), He is expressed in the three, eternally-existing *persons* of Father, Son, and Holy Spirit (a.k.a. the Triune God; Isaiah 40:1; 2nd Corinthians 13:14), each of Whom have distinct, yet inseparable operations in salvation (1st Peter 1:2). The second of the Triune, the Son, became a man (Luke 1:26-33; John 1:14), without losing His simultaneous position as God in the Triune (i.e. the Hypostatic Union; Colossians 2:9; John 10:30), without having a sin nature (Hebrews 4:15), and who committed no sin throughout His life (1st Peter 2:22). Jesus Christ experienced hardened resistance and persecution during His brief ministry (John 11:53; Luke 22:3-4), was put to death by crucifixion on a Roman cross (Matthew 27:33-46; John 19:34), and was buried in a guarded- and sealed- tomb (Matthew 27:62-66). Despite this, a bodily-resurrected Jesus Christ (Luke 24:36-43) exited the tomb after three days (Matthew 28:1-6; Mark 16:9-19), ministered for a short period of time to His disciples, then ascended into heaven in the presence of hundreds of witnesses (Acts 1:10-11; Hebrews 1:1-4), where He will remain until the time of His prophesied return (Acts 3:21; Jude 14-15).

Biblical Christianity, thus, exhorts that Jesus Christ's work on the cross paid a one-time substitutionary atonement for all entirety of believers' sins (Hebrews 9:12; 1:3; John 19:30) and Who also imputes onto these believers His perfect righteousness (2nd Corinthians 5:21). In essence, Christ's righteousness covers a believer's unrighteousness (Romans 3:10-18) with His own, reconciling the believer with God (Romans 5:1), averting eternal punishment (Romans 5:9), securing a co-inheritance of God's kingdom (Titus 3:7), and without the possibility of diminishing or losing one's "saved" status (Romans 8:1-2; John 6:37-40; John 10:27-28); that is: the believer with a true, saving faith is eternally assured of their salvation.

The process of this salvation starts with the electing to salvation by the Father (1st Peter 1:2-4; Ephesians 1:4), a drawing of the believer to faith by God (John 6:44), the drawn hearing and believing the Bible regarding Jesus Christ (Romans 10:17; Ephesians 1:13), confessing their belief of Jesus' resurrection and lordship (Romans 10:9-13; Matthew 10:32-33), and then repenting (i.e. confessing and turning away) from past and future sins (1st John 1:9-10). Biblical Christianity states that, while such salvation is certainly the product of a believer's action to believe (Ezekiel 18:32; Romans 3:5-6), it is primarily the act of faith, apart from works (Genesis 15:6, Habakkuk 2:4; Ephesians 2:8-9; Romans 4:4-5; Titus 3:5; 2nd Corinthians 3:5), that brings salvation, even though such faith *causally* produces good works (2nd Corinthians 5:17; 3:18; Philippians 2:13; Ephesians 2:10). Thus, under Biblical Christianity, a saving faith leads to justification which then causally leads to good works; not that good works, themselves or in combination with faith, are a causal agent of salvation from sins.

We've labored to describe as precisely and concisely as possible, the process of salvation under Biblical Christianity so that we can now discuss the superficially-subtle, yet eternally-consequential differences among the so-called Abrahamic faiths; using Biblical Christianity as a foundation and reference for comparison. We'll begin with Roman Catholicism (Reeves, 2010; Mortenson, 2015). First, Roman Catholicism derives its knowledge of and asserted authority not only from the sixty-six books of the Christian Bible but also from additions to the Old Testament by sources of questioned authenticity (e.g. Tobit) and inspiration (e.g. Maccabees I and II). Additionally, Roman Catholicism employs Papal Encyclicals, council pronouncements, Canon Law, and so on, carrying the same if not superior authoritative weight as the sixty-six books seen in Biblical Christianity. This immediately presents a difference between Roman Catholicism and Biblical Christianity in that the latter asserts that additions and subtractions should not be made (Proverbs 30:6; Deuteronomy 4:2; Revelation 22:19) to the completed work of the Bible (Jude 3; Luke 24:44), especially changes not directly inspired by the Holy Spirit (2nd Peter 1:20-21) or that include human traditions or worldly philosophies (Colossians 2:8).

What's more is that Roman Catholicism bases its authority, not only on its ever evolving body of spiritual understandings, but also from an ecclesiastical structure not directly defined in the Bible: the Roman Catholic Church and its human hierarchy, most prominent of whom is the Pope. Roman Catholicism claims their authority on their own interpretation of Scripture, most particularly Matthew 16:18-19, as well as other non-Biblical pronouncements originating from within the Roman Catholic Church itself. Biblical Christians refute this claim based on the charge of interpretive license as well as through the Bible itself which, Biblical Christianity asserts, denounces ecclesiastical elitism (1st Peter 5:2-4) and promotes Christ-centric church headship (Ephesians 1:22-23; 2:20-22; 1st Corinthians 3:11). What's clear is that both Biblical Christianity and Roman Catholicism differ markedly in Scriptural authority and permanence, with the former asserting a fixed, "Bible-only" view where the Bible itself is the ultimate

authority; the latter asserting a dynamic, "Bible-plus" view that may include human changes, dynamic interpretations, and human-deity authoritative equivalence.

These differences, we note, are not trivial matters; they directly shape the religious practices of both the church and its members. Roman Catholicism's Bible-plus perspective results in views that differ markedly from Biblical Christianity. For example, and under authority not found in the Bible itself, Roman Catholicism asserts that Jesus' mother, Mary, died a virgin, was without sin, and has taken a place of prominence within the heavenly order as a "co-redemptrix" or co-mediator with Christ. These beliefs are countered by Biblical Christians who note that the Bible unambiguously infers to Mary's loss of virginity (Galatians 1:19), her need of salvation and her subservience to the Lord Jesus (Luke 1:46-48), and that she nor anyone else is capable of serving in any mediatorial capacity other than Jesus Christ (John 14:6; Acts 4:12; 1st Timothy 2:5). These are two, very different belief systems which have their beliefs, worship, and religious practices at direct odds. Further, these are not points that can be easily addressed by ecumenical efforts, but are core beliefs that cannot be bridged by human endeavors.

Directly impacting the actions of Roman Catholics is Roman Catholicism's sacramental system that begins with infant baptism, believed to purge the recipient of original sin, and continues across time in the form of various church-mediated sacramental acts, the most prominent of which is the Mass. Here, Roman Catholics believe that the officiating priest is able to bring the literal person of Jesus Christ from heaven, transubstantiate (i.e. turn) ceremonial bread and wine into the literal body and blood of Christ, respectively, so that practicing Roman Catholic parishioners may partake in a periodic, quasi-sacrifice of Christ. This practice, along with other works of obedience and penance, is believed by Roman Catholics to reduce their time in a post-death limbo state of spiritual refinement called Purgatory; a concept believed by Roman Catholics since Catholicism's early days (e.g. Augustine) but formally accepted by the Roman Catholic Church with their inclusion of the uninspired Maccabees II into canon at the Council of Trent (1546 A.D.).

Biblical Christians strongly refute the need for a continual sacrifice of Christ's one-time atonement for sins (Hebrews 10:10-12), do not believe in the literal transformation of ceremonial bread and wine elements into the body and blood of Christ (John 6:47-54; Leviticus 17:12), and do not agree with the concept of a heaven-hell limbo state (i.e. Purgatory; Hebrews 9:27). This puts Biblical Christianity and Roman Catholicism at odds in that the former places unitary allegiance and obedience to Christ as noted in the Bible, whereas the latter places a combined allegiance and obedience to Christ and the Church. Further, as seen in this one example, this allegiance and obedience of the former is fixed and regulated only by the Bible, whereas the allegiance and obedience of the later is dynamic, evolving, subject to change, and delineated by more than the Bible alone.

Thus, from the discussion above, we see that Biblical Christianity and Roman Catholicism differ to a large degree on the ultimate authority of spiritual matters and what is required to obtain salvation. Biblical Christianity, in its most pure form, derives its authority from the Bible alone and an as-literal-as-possible interpretation thereof. Further, once saved by the one-time atoning work of Christ on the cross, Biblical Christians believe that additional good works are not a causal factor of maintaining salvation; salvation for the Biblical Christian can never be lost. Roman Catholicism, on the other hand, derives its authority from the Bible *and* other, man-originated pronouncements that may or may not fully align with the Bible in its original form or of the early Christian church. Further, practicing Roman Catholics are responsible for maintaining obedience to both Christ and the Roman Catholic Church on penalty

of post-death punishment in Purgatory or even damnation in the lake of fire; salvation for the Roman Catholic is never assured in this lifetime and can even be lost.

All of these differences, then, will necessarily translate into how investment managers allocate clients' wealth within Christian-centric or Catholicism-centric investment funds; the former being based on a fixed, literal interpretation of the Bible and the latter being based on a dynamic, set of Biblical interpretations and human-centric edicts. Similarly, Biblically Christian- and Roman Catholic- investors will necessarily have their own, faith-specific investment utility functions that will only partially and imperfectly overlap with one another: demonstrably different faiths with clearly different bases for salvation, obedience, compliance, and so on, will almost assuredly translate those beliefs into different investment screening practices. In short, Biblically Christian- and Roman Catholic- investment participants should differ, the funds they participate in should differ, and these differences should be reflected in the performance, attribution, and style characteristics of those investments.

Thus, the seemingly-equivalent Biblically Christian and Roman Catholic religions that ecumenicists are so quick to call "brother faiths" are, clearly, very different in beliefs, practices, and outcomes. This stark difference between Biblical Christianity and Roman Catholicism is, relatively speaking minor, compared to the unbridgeable differences in beliefs between Christ-centric faiths and the Islamic tradition (Caner, 2015). For example, Islam refutes the deity and atoning work of Christ, instead asserting that Christ was only a prophet and messenger of Allah. In Christianity, there is no salvation without Christ, meaning that Christianity and Islam differ on salvation, right out of the gate. Even the concept of deity and the Godhead is different between Christianity and Islam. Biblical Christianity and Roman Catholicism rest on a shared concept of a time-and-creation independent Triune God (i.e. one monotheistic *essence* expressed in three *persons*); monotheistic in *essence* but separated in *personhood*. Islam, equally asserts a monotheistic deity, but that this deity is monomodal or unitarian in nature; there is no distinction within the godhead. Hence, and by consequence, while the Islamic deity is capable of transcendence relative to humanity, it is not transcendent in a fully, self-sufficient, creation- and time- independent manner; the Islamic deity is constantly-evolving as its interaction with creation co-shapes the two over time.

Further, Islam utilizes a very different source of information than the Bible. Biblical Christianity, for example, insists that its spiritual understanding comes explicitly from the work of the Holy Spirit (i.e. God) inspiring Old- and New- Testament authors to write God's words. Islam, on the other hand, derives its spiritual understanding from the recorded oral traditions of the Prophet Muhammad's various proclamations throughout his lifetime, most important of which is his reported interactions with Allah's messenger angel, Gabriel. The Qur'an was finalized some 500 years after the early Christian church had completed its canon and, in similar fashion to the Roman Catholic canon, has been evolving over time. Further, despite Biblical Christians' insistence of their Bible's unchanging validity over time and translations (i.e. transmissible fidelity), and Islam's own use of Biblical themes within the Qur'an, Islam asserts that both the Old and New Testament were intentionally corrupted and, thus, invalid for representing spiritual truth in their current forms. Furthermore, while Biblical Christianity, and to a much lesser degree Roman Catholicism, asserts that a believer's salvation is assured throughout all time and cannot be lost, Islam's concept of salvation has no "resting point" during one's lifetime; its is only during a post-death judgment that the Islamic believer discovers their eternal resting place. And, finally and consequently, as the Islamic deity is evolving over time, as is the

body of understanding within and around the Qur'an, Islamic believers' adherence to a lifetime of obedience to the Qur'an may similarly evolve.

Thus, while again this paper is not intended to be an authoritative reference on faith traditions or comparative apologetics, we can see from the detailed discussion above, that Biblical Christianity, Roman Catholicism, and Islam are all three very different religions. Consequently, we should not be surprised if the investing behavior of faith-based managers, and the investment demand of faith-based investors, should differ markedly among the oftentimes erroneously homogenized faiths. We believe that the empirical results presented within this study demonstrate exactly that.

Macroeconomic Determinants of the Credit Loss Forecasting

Zilong Liu*, Hongyan Liang†, Chang Liu‡, and Yang Cheng§

Abstract

Macroeconomic variables are critical inputs in credit loss forecasting (LF) models and are mandated by regulators for stress testing to project potential credit losses under various economic scenarios. The COVID-19 pandemic introduces unprecedented volatility to macroeconomic indicators, disrupting their historically strong relationship with credit loss rates and challenging the robustness of LF and Current Expected Credit Loss (CECL) models.

This study examines the dynamic relationship between charge-off rates for U.S. commercial banks' loan portfolios—commercial and industrial (C&I) loans, consumer loans, and real estate loans—and macroeconomic factors from 1990 Q1 to 2024 Q3. By employing advanced machine learning and traditional regression-based approaches, we identify key macroeconomic variables, such as labor market indicators, housing market dynamics, and consumer financial conditions, that drive credit risk across these loan categories. The results highlight the need to reevaluate traditional relationships to ensure model robustness, particularly in the context of COVID-19-era data. This study provides practical insights for banks to enhance credit risk modeling frameworks, offers guidance for macroeconomic variable selection in LF and CECL applications, and supports policymakers in designing effective regulatory frameworks during periods of economic uncertainty.

Keywords: Macroeconomic factors; Loss forecasting; COVID-19 recession; Stress testing; CECL; Credit risk

JEL classification: C32, E17, G20

I. Introduction

In the past several decades, extensive research is conducted in the risk management area, especially in credit risk management. Credit risk is the most common risk for banks and is closely monitored by regulators. The recent global financial crisis (GFC) heightens the importance of credit risk management to financial institutions. After the GFC, bank regulators tighten the supervision in financial industry under the Dodd-Frank Act Stress Tests (DFAST). In particular, large banks are required to calculate their capital ratio under a hypothetical stressed economic scenario to determine their insolvency risk under a negative economic shock. In practice, banks have to incorporate scenario analysis and forward-looking macroeconomic indicators more rigorously, which leads to the evolution of credit-loss forecasting techniques. Therefore, the understanding of the relationship between macroeconomic factors and charge-off rates on loans becomes a worldwide issue for banks, policymakers, and regulators.

* Gies College of Business, University of Illinois Urbana-Champaign, Email: zilongl@illinois.edu

† Gies College of Business, University of Illinois Urbana-Champaign, Email: liangh@illinois.edu

‡ Harmon College of Business and Professional Studies, University of Central Missouri, Email: cliu@ucmo.edu

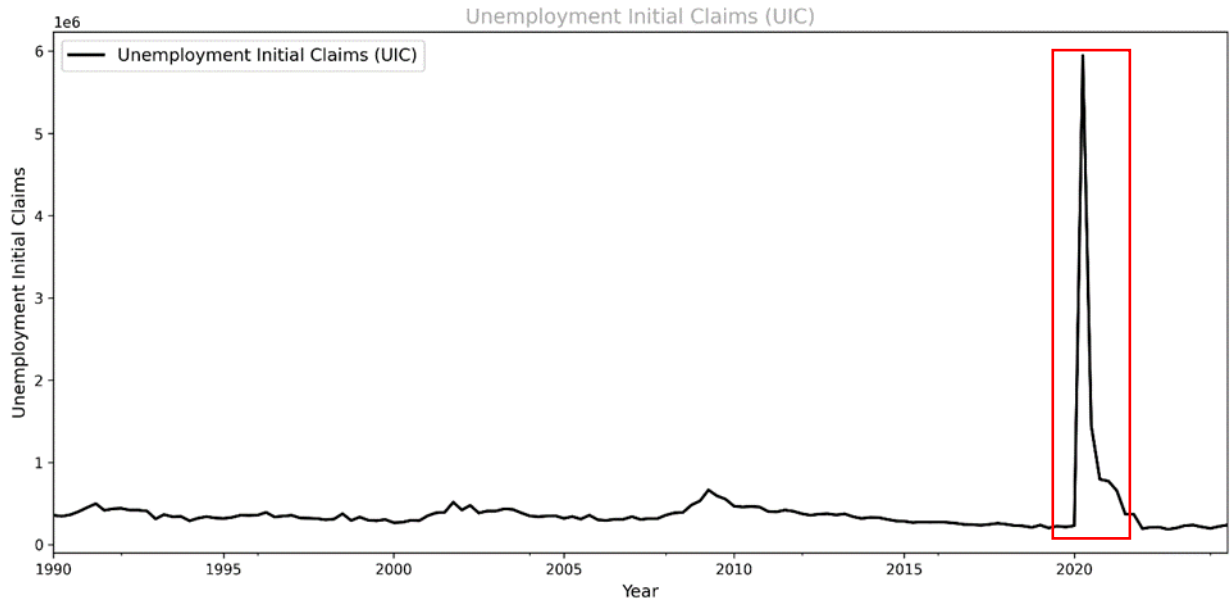
§ Craig School of Business, College of Business and Professional Studies, Missouri Western State University, Email: ycheng@missouriwestern.edu

The stress testing requirement on banks initiates debates on a number of issues in the literature on credit risk management. The first issue is whether the macroeconomic environment and business cycles play a vital role in determining the delinquency risk of loans. There is a large body of literature study the link between charge-off risk and the overall economic climate (Virolainen 2004; Mählmann 2005; Hackbarth et al. 2006; Pesaran et al. 2006; Jakubik 2007; Bonfim 2009; Castro 2013; Djeundje and Crook 2018; Breeden and Crook 2020). They find that macroeconomic conditions have a direct impact on bad debts and charge-offs. In addition, it is also widely recognized that financial crises are more likely to occur in adverse economic conditions, where the unemployment rate increases and the gross domestic product (GDP) growth rate declines. For example, Bonfim (2009) finds that macroeconomic variables have a significant contribution to credit loss after controlling firm-level characteristics, suggesting that macroeconomic conditions are critical when assessing the firm's credit risk.

Starting in March 2020, the COVID-19 pandemic led to a "shelter in place" order across the country, which significantly disrupted economic activities. This disruption resulted in significant job losses and deterioration in the economic outlook. At the same time, both the U.S. government and financial institutions deployed several measures to help residents remain afloat during these uncertain times. The sudden change in the macroeconomic environment, as well as the government invention, led to extremely high volatility in the macro-variables. For example, the variables in the labor market category deteriorated significantly at the onset of the pandemic. As a result of the entire country's lockdown, the unemployment initial claims skyrocketed by more than 20 times from about 1 million in March 2020 to almost 20 million in April 2020 (Figure 1). However, the personal income variables experienced significant improvements (increases) because of the government stimulus bill (Figure 2). Unlike the extreme volatility observed in the macroeconomics variables, the charge-off rates on loans at U.S. commercial banks didn't increase. For example, the delinquency rate and charge-off rate for commercial and industrial (C&I) loans, consumer loans, and real estate loans dropped from Mar 2020 to Dec 2020 as shown in Figure 3. The unresponsive loss rates under pandemic were a result of various government relief supports and corporate forbearance programs. The volatile macroeconomic variable movement and delayed loss recognition process together contributed to the performance deterioration of loss forecasting/CECL models.

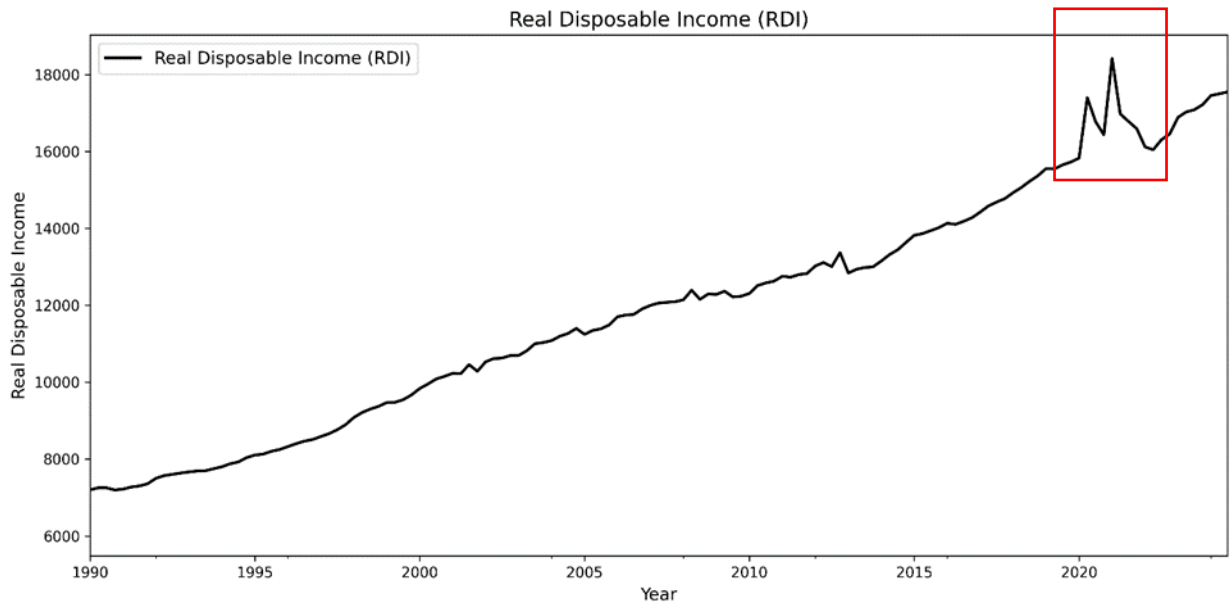
Although prior studies identify macroeconomic variables as key determinants of credit risk, the historically observed strong relationships between macroeconomic variables and charge-off rates disappeared during the COVID-19 pandemic. The volatile movement of macroeconomic variables, combined with the delayed loss recognition process, leads new challenges for banks when they use macroeconomic variables to model credit risk. There are two concerns for banks. First, whether the prior observed relationships between macroeconomic factors and loan charge off rates are still applicable to the current situation. Second, which macroeconomic variables have the strongest predictive power for future loan default rates given the COVID-19 impacts. Our study will shed light on these two questions.

Figure 1: Unemployment Initial Claims

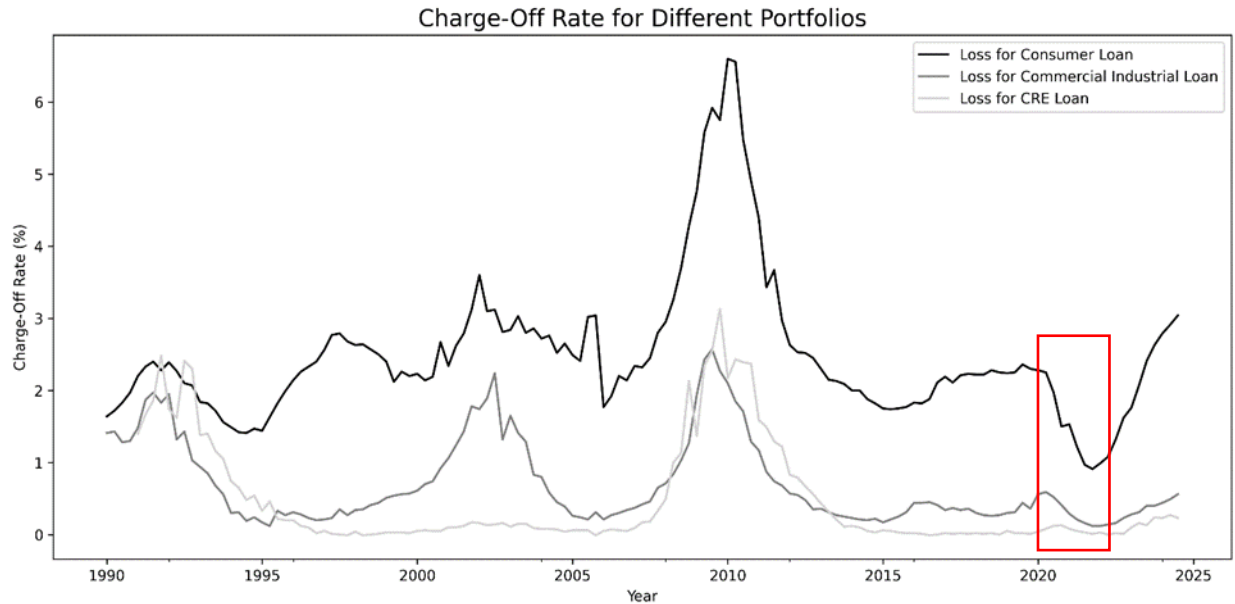


Source: U.S. Department of Labor

Figure 2: Real Disposable Personal Income



Source: U.S. Bureau of Economic Analysis

Figure 3: Charge-Off Rate for Loans of all U.S. Commercial Banks

Source: Federal Reserve Board

Most of the prior research is conducted based on data from the pre-pandemic period. Therefore, those studies have data limitations that are unable to capture the impact of the COVID-19 pandemic. Along with the fact that the recent data shows a missing link between the key macroeconomic variables and charge-off rates on loans, it is unclear whether the prior observed relationships between macroeconomic factors and default rates on loans still hold after including the pandemic data. The Federal Reserve heightens the standard on bank monitoring and highlights the importance of exploring a more robust relationship between macroeconomic variables and credit loss that can be used during the interim of the COVID-19 pandemic. Given the uncertainty about the economic outlook is quite large under COVID-19, ensuring the stability and accuracy of the loss forecasting framework are imperative for banks, regulators, and policymakers. This study helps address these concerns and our research will fill this gap by incorporating the data of the COVID-19 pandemic period. To the best of our knowledge, this is the first paper to study the relationship between macroeconomic factors and loan delinquency including the COVID-19 impacts. The results present in our research have practical implications for the financial industry and can be directly used by banks to build or enhance their credit loss/reserve forecast models. Finally, the findings in our research can help regulators and policymakers to design appropriate guidelines for the use of advanced machine learning models in the banking industry.

This paper is structured as follows. In the following section, we review the related literature. In Sections III and IV, we describe the data sources and present the descriptive statistics of all variables used in this study. In Sections V and VI, we move to present our analysis and empirical evidence. We present model selection and prediction in section VII. Finally, we discuss the results and summarize the conclusion in Section VIII.

II. Literature Review

The impact of macroeconomic variables (i.e., GDP growth, unemployment rates, and inflation) and banking sector variables (i.e., loan-to-deposit ratios and capital adequacy) on credit risk and non-performing loans are discussed extensively in the prior literature as they influence both borrower behavior and bank resilience. Agarwal and Liu (2003) find that country-level unemployment rates significantly impact credit card delinquency rates, highlighting the connection between regional macroeconomic fluctuations and delinquency rates. To capture non-linear relationships between macroeconomic conditions and defaults, Pesaran et al. (2006) introduce an innovative approach to modeling conditional credit loss distributions by isolating the systematic risk from firms' idiosyncratic risk. Their findings reveal that the default probabilities are linked to business cycles. Giesecke et al. (2011) further investigate the predictive power of macroeconomic variables for U.S. corporate bond defaults using a regime-switching model. Their results indicate that stock returns, stock return volatility, and GDP changes are robust predictors of default rates. These predictors reflect that both market dynamics and economic health are critical in assessing creditworthiness. Furthermore, macroeconomic factors not only affect the default rates but also influence the credit loss resulting from non-performing loans. Betz et al. (2020) find that loan resolution time doubled during a recessionary period compared to benign periods due to higher borrower defaults and strained banking resources. This finding indicates that the prolonged resolution process results in highly skewed loss distributions and elevates levels of systematic risk within bank portfolios.

The relationship between macroeconomic factors and credit risk are widely documented beyond the United States. Fung and Wong (2002) examine interaction of macroeconomic variables, bankruptcies, and credit card charge-off ratios in Hong Kong. Their findings reveal that the CPI, unemployment rate, bankruptcy cases, and credit card charge-off percentages are cointegrated, suggesting a long-term equilibrium relationship among these variables. Declines in the CPI are associated with falling asset values and reduced economic activity, which contribute to rising unemployment, bankruptcies, and personal credit defaults. These findings explain how commercial banks predict the risk of retail credit operations and estimate credit loss provisioning. Similarly, Castro (2013) analyzes data from five European countries, Greece, Ireland, Portugal, Spain, and Italy (GIPSI) from 1997 to 2011 and concludes that the macroeconomic conditions significantly influence banking credit risk. Specifically, credit risk tends to increase during periods of declining GDP growth, rising unemployment, and falling house prices. Additionally, real exchange rate appreciation is found to have a positive correlation with credit risk. Koju et al. (2019) extend the scope of analysis by examining the effects of macroeconomic variables on credit risk across 49 developed countries between 2000 to 2015. Their findings indicate that loan performance and default risk are closely related to the growth of industrial sectors and export activities. The study also highlights that expansionary fiscal policies can play a crucial role in strengthening the stability of the banking system.

Although prior research suggests that macroeconomic conditions are linked to credit risk, the current literature could not find consistent evidence for the effects of macroeconomic factors on charge-offs over different business cycles. Ausubel (1997) finds that loan default rates rise during a benign economic environment in which GDP is growing moderately and unemployment is low. This finding seems against the foundational belief that the charge-off rate will decrease in robust economic periods and increase in a downturn period. More recent research shows that other economic factors may contribute to charge-offs. Stavins (2000) identifies that personal debt

variables, such as debt-to-disposable income ratio, have a strong correlation with credit card charge-offs and bankruptcy rates. Gross and Souleles (2002) examine the relationship between personal bankruptcy and credit card defaults, concluding that the impact of macroeconomic factors, such as unemployment rate, on charge-off rate changes significantly over time and that there is no conclusive evidence to prove a relationship between charge-off rate and macroeconomic factors.

The existing literature also points out that macroeconomic factors alone cannot perfectly predict the future default rate, they have to be used in conjunction with other factors, such as firm-specific factors or unobservable factors, to predict the default rates. Pu and Zhao (2012) examine a comprehensive set of credit risk drivers, including industry and macroeconomic factors. They find that there is an economically significant credit risk that remains to be unexplained even after controlling observable factors. They also suggest that unobservable risk factors should be incorporated into credit risk models in addition to macroeconomic factors. Figlewski et al. (2012) study how different factors, including overall economic conditions and specific company characteristics, affect loan losses (charge-offs) and big changes in credit ratings. They find that both factor categories are significant, but macroeconomic variables are highly dependent on the inclusion of other factors. In an extension of their work by Bellotti and Crook (2013), a discrete-time survival model is proposed to predict the probability of charge-off. They claim that using macroeconomic variables along with behavioral factors produces the best predictive fit.

Our research is also related to another strand of literature on using macroeconomic factors to forecast future credit losses. Liu and Xu (2003) adopts a step-wise regression and vector autoregression to select macroeconomic variables that are useful for predicting credit card charge-offs in the U.S. By using the data from the period of 1986-1998. They find the unemployment rate, consumer confidence index, household debt service burden, inflation rate, personal bankruptcy filings, and stock market returns are powerful variables in predicting the future charge-off rate. However, their research has limitations. The data period is relatively short and excludes the GFC period. In addition, their sample period also does not reflect the most recent changes in the credit card industry. Taghiyeh et al. (2021) adopts more advanced model selection techniques and machine learning algorithms to build models that can be used to forecast credit card charge-off rates. They use 19 key macroeconomic indicators as the potential drivers and find that unemployment initial claims have the highest relative importance in determining the credit loss forecast.

Prior studies, such as those by Liu and Xu (2003) and Taghiyeh et al. (2021), extensively analyzes the macroeconomic determinants of credit risk; however, they can't capture the unprecedented economic disruptions caused by the pandemic. Our study uniquely extends this line of inquiry by analyzing how these extraordinary circumstances alters previously observed economic relationships, providing new insights into credit risk modeling during periods of severe economic stress.

III. Data

This study uses quarterly observations of charge-off rates on commercial and industrial loans, consumer loans, and real estate loans of all commercial banks from 1990 Q1 to 2024 Q3. The loan charge-off rate data are downloaded from the Board of Governors of the Federal Reserve System (U.S.). Although the credit charge-off loss data begins in 1985, some key macroeconomic indicators (i.e., market volatility index) only begin to available in 1990 Q1. We believe the data period of 1990 Q1 to 2024 Q3 is sufficient long because it covers multiple business cycles

including 1990 oil price shock, 2000 internet bubble, 2007-2008 global financial crisis as well as the COVID-19 pandemic period. The historical macroeconomics variables are downloaded from the Board of Governors of the Federal Reserve System. The Federal Reserve consolidates the historical data as well as forecasting series of 28 core macroeconomic indicators (14 domestic and 14 international) for bank stress testing purposes. The 28 variables cover the key indicators from different categories of macroeconomic variables (i.e., unemployment rate represent the condition in the labor market). Because we are interested in the charge-off rates in the U.S., only the 14 domestic macroeconomic variables are used in this study. In addition, we also obtain some key macroeconomic indicators not provided by the Federal Reserve, such as non-farm employment, retail sales, and personal bankruptcy, from the Bureau of Labor Statistics and the U.S. Bureau of Economic Analysis. It is worth noting that Household Financial Obligations (HFO) data is only available until 2023 Q4, resulting in a total of 135 observations for this variable. The total number of macroeconomic variables used in the study is 26, in addition to the 3 charge-off rates. Detailed descriptions of the data and their sources are provided in Appendix A.

For selected macroeconomics variables, we also consider the following transformations: i) logarithm transformation; ii) year-over-year (yoy) change, iii) quarter-over-quarter (qoq) change, and iv) 1 to 4 quarters lagged of both raw and transformed variables. The yoy and qoq changes are calculated by using the logarithm growth rate instead percentage growth rate because the logarithm growth is more symmetric and less skewed than the percentage growth rate.

IV. Descriptive Statistics

Table 1 reports descriptive statistics for all variables used in this study. The average charge-off rates of real estate loans, consumer loans, and C&I loans over the sample period are 2.48%, 0.70%, and 0.47%. On average, unemployment rate is 5.76%. In addition, the average mortgage rate, prime rate, 3-month treasury rate, and 5-year treasury rate are 6.01%, 5.86%, 2.66%, and 3.76%, respectively.

Table 2 reports the pair-wise correlation coefficients of the variables. To save space, we only report the correlation coefficients of macro factors and charge-offs rate. The correlation coefficients among the charge-off rates of three different loans are in the range of 0.62 to 0.71, indicating there is some heterogeneity across products. For C&I loans, unemployment rate (UR) and BBB corporate yield have the highest correlation coefficients; for consumer loans, UR and debt service burden are highly correlated with the charge-off rate; for real estate loans, UR, new housing unit, and non-farm employment exhibit the highest correlation. These results suggest that UR is a shared potential risk factor for all three loans. In addition, the correlation coefficients for the same macro variables show significant differences across loans, suggesting the risk drivers for different loans are different. For example, the correlation coefficient of BBB corporate yield with consumer loans charge-off rate is only 0.18 while with C&I loans charge-off rate is 0.46. Therefore, it is reasonable to separate the analysis for different loans in our study.

Table 1 Descriptive Statistics

This table presents the descriptive statistics for the charge-off rates and macroeconomic variables used in the study. The data spans from 1990 Q1 to 2024 Q3, covering multiple business cycles. We include variable distributions and their respective sources for GDP growth, unemployment rate, interest rates, and other financial metrics.

Variable Name	N	Mean	Std. Dev.	Min	Median	Max	Source
Real Estate Loans Charge-off Rate	139	2.48	0.998	0.91	2.28	6.60	Federal Reserve Board
Consumer Loans Charge-off Rate	139	0.70	0.588	0.12	0.44	2.57	Federal Reserve Board
C&I Loans Charge-off Rate	135	0.47	0.739	-0.01	0.10	3.13	Federal Reserve Board
Real GDP	139	16101.93	3751.73	9951.92	16420.74	23400.29	U.S. Bureau of Economic Analysis
Nominal GDP	139	14463.34	6190.37	5872.70	14381.24	29374.91	U.S. Bureau of Economic Analysis
Real Disposable Income	139	11882.53	3050.02	7195.15	12055.32	18411.66	U.S. Bureau of Economic Analysis
Nominal disposable Income	139	10743.84	4745.07	4232.90	10406.08	21710.15	U.S. Bureau of Economic Analysis
UR	139	5.76	1.82	3.40	5.40	14.80	U.S. Bureau of Labor Statistics
CPI Inflation	139	205.83	48.10	127.50	205.90	313.53	U.S. Bureau of Labor Statistics
3M Treasury Rate	139	2.66	2.24	0.02	2.33	7.77	Federal Reserve Board
5Y Treasury Rate	139	3.76	2.16	0.28	3.69	8.77	Federal Reserve Board
10Y Treasury Rate	139	4.27	1.99	0.62	4.15	8.79	Federal Reserve Board
BBB Corporate Yield	139	5.93	2.05	2.06	5.83	10.90	WARGA and Merrill Lynch Database
Mortgage Rate	139	6.01	1.87	2.67	6.16	10.22	Federal Home Loan
Prime Rate	139	5.86	2.28	3.25	5.50	10.50	Federal Reserve Board
Dow Jones Index	139	13914.27	9437.37	2452.48	10850.66	39807.37	Dow Jones
House Price Index	139	327.66	129.71	165.25	323.26	688.42	CoreLogic
CRE Price Index	139	191152.80	78162.63	86351.00	179529.00	351613.00	Federal Reserve Board
Market Volatility Index	138	19.59	7.69	9.51	17.67	53.54	Chicago Board Options Exchange
Non-Farm Employment	139	132816.50	13115.42	108290.00	132455.00	158692.00	U.S. Bureau of Labor Statistics
Retail Sales- Motor	139	431.32	132.10	120.15	449.00	658.90	U.S. Bureau of Economic Analysis
New Housing Unit	139	1351.99	406.20	521.00	1378.00	2212.00	U.S. Dept. of Housing and Urban Dev.
Personal Consumption Expenditures (PCE)	139	9669.40	4249.90	3730.70	9643.60	19866.30	U.S. Bureau of Economic Analysis
Personal Saving	139	6.15	3.31	1.40	5.80	32.00	U.S. Bureau of Economic Analysis
WTI Oil Price	139	51.10	29.70	12.47	47.22	133.44	Federal Reserve Bank of St. Louis
Initial Claims	139	396230.20	493968.70	187000.00	340000.00	5946000.00	U.S. Employment and Training Admin.
PPI	139	168.34	42.69	114.10	169.10	272.27	U.S. Bureau of Labor Statistics
Debt Service Burden	139	12.09	1.56	9.08	11.65	15.85	Federal Reserve Board
Household Financial Obligations	135	16.07	1.20	12.43	16.32	18.16	Federal Reserve Board

Table 2 Correlation Matrix

Table 2 presents the correlation matrix for the charge-off rates and macro variables. Column (1) presents the correlations between real estate loans and macro variables. Column (2) presents the correlations between consumer loans and macro variables. Column (3) presents the correlations between C&I loans and macro variables.

Variables	(1)	(2)	(3)
(1) Real Estate Loans Charge-off Rate	1.00		
(2) Consumer Loans Charge-off Rate	0.62	1.00	
(3) C&I Loans Charge-off Rate	0.70	0.71	1.00
(4) Real GDP Growth	-0.32	-0.07	-0.32
(5) Nominal GDP Growth	-0.28	-0.13	-0.33
(6) Real Disposable Income	-0.28	-0.07	-0.29
(7) Nominal Disposable Income	-0.26	-0.12	-0.31
(8) UR	0.66	0.49	0.46
(9) CPI Inflation	-0.25	-0.06	-0.30
(10) 3M Treasury Rate	-0.14	-0.24	-0.10
(11) 5Y Treasury Rate	0.07	-0.13	0.10
(12) 10Y Treasury Rate	0.20	-0.04	0.21
(13) BBB Corporate Yield	0.35	0.18	0.46
(14) Mortgage Rate	0.19	-0.03	0.24
(15) Prime Rate	-0.14	-0.21	-0.08
(16) Dow Jones Index	-0.40	-0.29	-0.40
(17) House Price Index	-0.28	-0.12	-0.29
(18) CRE Price Index	-0.35	-0.19	-0.35
(19) Market Volatility Index	0.18	0.36	0.40
(20) Non-Farm Employment	-0.47	-0.09	-0.35
(21) Retail Sales- Motor	-0.07	-0.15	0.00
(22) New Housing Unit	-0.69	-0.45	-0.37
(23) PCE	-0.27	-0.11	-0.32
(24) Personal Saving	0.13	-0.12	0.05
(25) WTI Oil Price	0.09	0.18	-0.14
(26) Initial Claims	0.05	0.06	0.10
(27) PPI	-0.14	-0.06	-0.31
(28) Debt Service Burden	0.24	0.61	0.31
(29) Household Financial Obligations	0.17	0.43	0.42

V. Correlation Analyses

Correlation without Optimal Transformation

This section contains the results for correlation analysis. All macro variables have several transformations to capture the lag-lead relationships between the charge off rates and macro variables. The initial macro variables list only contains 26 unique variables (listed in Table 2), while after transformation there are 266 variables in total.

We conduct two types of correlation analysis — correlation without optimal transformation and with optimal transformation. In the without optimal transformation analysis, we first compute the correlation coefficients between the charge off rates and each transformed variable, then we rank transformed variables by the correlation coefficients in descending order and count how many

times a macro variable is selected among the top 50 list. The more frequently a variable is selected in the top 50 list, the more likely this variable is highly correlated with credit loss. For a more comprehensive review, we report the ranking results using three different correlation measures—Pearson, Spearman, and Kendall Tau correlation coefficients. Pearson correlation coefficient is widely used in the research as it measures the linear relationship among variables. In contrast, the Spearman correlation coefficient can capture the nonlinearity relationship among variables. Finally, Kendall's Tau rank correlation is a nonparametric measure and not affected by nonlinearities because Kendall's Tau relies only on the concordance of the two variables. This correlation measure is used in prior research of credit risk. The results for the correlation without optimal transformation are displayed in Table 3.

Table 3 Correlation without optimal transformation

This table presents the frequency of macroeconomic variables being selected in the top 50 correlations with charge-off rates across three measures—Pearson, Spearman, and Kendall's Tau—for C&I loans, consumer loans, and real estate loans in Panel A, Panel B, and Panel C, respectively.

Panel A: C&I Loans - Number of times selected in the Top 50 list based on

Variable Name	Pearson	Spearman	Kendall's Tau	Row Total
Unemployment Rate	3	10	10	23
Non-Farm Employment	3	8	9	20
3M Treasury Rate	6	4	4	14
BBB Corporate Yield	5	4	4	13
House Fin. Obligations	4	4	4	12
CRE Index	4	3	3	10
Market Volatility Index	3	3	2	8
New House Unit	3	2	2	7
Initial Claims	0	3	3	6
5 Year Treasury Rate	0	3	3	6
Real GDP Growth	4	0	0	4
Mortgage Rate	1	2	1	4
Nominal GDP Growth	3	0	0	3
House Price Index	3	0	0	3
DJ Index	1	0	0	1
Grand Total	50	50	50	150

Panel B: Consumer Loans - Number of times selected in the Top 50 list based on

Variable Name	Pearson	Spearman	Kendall's Tau	Row Total
Non-Farm Employment	3	10	10	23
Unemployment Rate	2	10	10	22
House Fin. Obligations	6	7	7	20
Debt Service Burden	6	6	7	19
New House Unit	6	3	3	12
House Price Index	10	0	0	10

Market Volatility Index	0	5	4	9
Nominal GDP Growth	5	1	1	7
Initial Claims	0	3	3	6
Personal Saving	0	3	3	6
CRE Index	5	0	0	5
Real GDP Growth	4	0	0	4
DJ Index	2	0	0	2
Consumption Exp.	1	0	0	1
3M Treasury Rate	0	1	0	1
5 Year Treasury Rate	0	1	0	1
BBB Corporate Yield	0	0	1	1
Prime Rate	0	0	1	1
Grand Total	50	50	50	150

Panel C: Real Estate Loans - Number of times selected in the Top 50 list based on

Variable Name	Pearson	Spearman	Kendall's Tau	Row Total
House Price Index	10	10	10	30
Non-Farm Employment	3	11	11	25
CRE Index	9	5	5	19
New House Unit	8	5	5	18
Unemployment Rate	4	6	7	17
Initial Claims	0	5	5	10
BBB Corporate Yield	3	3	3	9
Real GDP Growth	3	3	2	8
DJ Index	1	2	2	5
Prime Rate	4	0	0	4
3M Treasury Rate	3	0	0	3
House Fin. Obligations	1	0	0	1
Debt Service Burden	1	0	0	1
Grand Total	50	50	50	150

The results in Table 3 show that, UR, non-farm employment, 3-month treasury rate, and BBB corporate yield have a higher frequency of being selected among all the variables across the three correlation measures for C&I loans. For consumer loans, non-farm employment, UR, and house financial obligations have a higher frequency of being selected across all three correlation measures. Finally, house price index, non-farm employment, and CRE index are more likely to be selected in the real estate loans. It is worth noting that non-farm employment is highly rated by Spearman and Kendall's Tau correlation coefficients while ranked relatively low in the Pearson measure in both consumer loans and real estate loans. These results imply that non-farm employment is more concordant with the consumer and real estate charge-off rates and also have a strong nonlinear relationship.

Correlation with Optimal Transformation

Correlation with optimal transformation analysis only allows each macro variable to show up once in the top list by using its optimal transformation. The advantage of this analysis is that we can observe the optimal transformation as well as the magnitude of the correlation for each variable. In this section, we only select one transformed value that has the highest correlation coefficient with credit loss within each macro variable and its transformations, and then the transformed macroeconomic variables are ranked by correlation coefficients in descending order.

Table 4 displays the variables with their optimal transformation being selected based on three different correlation measures. We rank the variables from highest to lowest based on its correlation coefficients. The optimal transformation used and correlation coefficients are reported in Table 4. The p-value for all the transformed variables are 0.00. Therefore, they are omitted in the Table 4.

The results In Table 4 show that there are some common risk factors, such as labor market variables and housing market variables, across the three loan categories. However, the importance of these factors varies by loan type. For C&I loans, labor market variables, such as the unemployment rate and non-farm employment, are highly correlated with charge-offs, along with interest rate variables like the 3-month Treasury rate and the prime rate. For consumer loans, household financial conditions, including the debt service ratio and household financial obligations, show the strongest correlations, followed by variables related to housing, such as the house price index and new home units. Additionally, personal financial variables, such as personal savings and nominal GDP, are also significant. For real estate loans, housing market variables, including new home sales and the house price index, are dominant, with unemployment rate variables playing a secondary role. Across all categories, the transformations (e.g., year-over-year, lagged, or logarithmic) highlight the different dynamics and time sensitivities of these macroeconomic factors in explaining credit losses.

VI. Feature Selection

The correlation analysis examines which variables are most strongly associated with credit loss by evaluating the strength and direction of individual relationships. The correlation analysis is univariate, focusing on one variable at a time to determine its standalone correlation with credit loss. In contrast, feature selection is multivariate, considering the combined effects and relative importance of variables when used together in a predictive model. While correlation analysis helps identify promising candidates for inclusion, feature selection determines which variables contribute most significantly to the model when considered in conjunction with others. By combining insights from both approaches, we can better understand the unique and collective importance of variables in predicting credit loss.

To examine the variable importance across different macroeconomic variables, we use four feature selection algorithms – stepwise selection, gradient boost machine (GBM), random forest (RF), and Lasso regression. The rationale behind using four different feature selection algorithms is to include both black-box (with less interpretability) and regression-based models with a high level of interpretability. GBM and RF are black-box models which can capture complex nonlinear trend, consequently improving forecasting accuracy. The traditional stepwise regression has high interpretability. However, it only focuses on selecting between the correlated covariates without

considering the improvement in prediction accuracy. Finally, the lasso penalty regression not only improves the forecasting accuracy but also mitigates the overfitting issue.

In order to avoid the abuse and misuse of machine learning algorithms as well as the circles of “Garbage in, garbage out”, we further narrow down the number of variables used in the feature selection algorithms based on the results from the correlation analysis. For the feature selection stage, we include only the top 10 transformed variables that have the highest Pearson correlation coefficients. The final result is reported in Tables 5, 6, and 7 for C&I loans, consumers loans, and real estate loans, respectively. We retain the top 5 most important variables for each algorithm based on their relative importance.

In Panels A of Tables 5 to 7, we report the variables selected by stepwise and lasso regressions. Both models are generalized linear models. The variables selected by them show similarities, with lasso regression typically including more variables due to its penalty mechanism. In Panels B, the variables selected by the two tree-based models, GBM and RF, are presented. These black-box models capture nonlinear relationships. While the variables they select are largely consistent, RF tends to include a broader set of variables.

For the C&I loans charge-off rate, non-farm employment is the most important variable in tree-based models. Variables such as the CRE price index and new home units also demonstrate strong predictive power, emphasizing the importance of labor market and commercial real estate dynamics for C&I loan performance.

For the consumer loans charge-off rate, Panel B of Table 6 highlights that the house price index is the most important variable across tree-based models. In addition, variables reflecting household financial conditions, such as the debt service burden, and labor market indicators like the unemployment rate, play significant roles. Macroeconomic indicators, including GDP growth and consumer expenditure, also contribute meaningfully to the prediction of consumer loan charge-offs.

Finally, for the real estate loans charge-off rate, house price index has the highest predictive power across all models, followed closely by new home units and other housing market variables such as the CRE price index. These findings proof the dominant role of housing market variables in determining the performance of real estate loans.

Overall, the results suggest that while certain variables, such as labor market indicators, play a crucial role across multiple loan types, the relative importance of housing market and consumer financial variables varies depending on the loan category. The combination of feature selection algorithms highlights the diverse set of factors influencing credit losses across different loan products.

Table 4 Correlation with Optimal Transformation (Ranked from High to Low)

This table ranks macroeconomic variables by their highest correlation coefficients with charge-off rates across three measures—Pearson, Spearman, and Kendall’s Tau—using their optimal transformations for each loan category.

Panel A: C&I Loans

Variables	Transformation	Pearson	Variables	Transformation	Spearman	Variables	Transformation	Kendall
Unemployment Rate	yoy_lag1	61.63%	Non-Farm Employment	yoy	-71.54%	Unemployment Rate	yoy_lag1	55.59%
Non-Farm Employment	yoy	-54.14%	Unemployment Rate	yoy	71.49%	Non-Farm Employment	yoy	-55.08%
3-Month Treasury Rate	yoy	-58.83%	3-Month Treasury Rate	yoy	-60.97%	Prime Rate	yoy	-43.67%
Prime Rate	yoy	-61.13%	Prime Rate	yoy	-59.79%	BBB Corporate Bond Yield	lag4	43.48%
BBB Corporate Bond Yield	lag4	50.35%	BBB Corporate Bond Yield	lag4	58.13%	3-Month Treasury Rate	yoy	-42.77%

Panel B: Consumer Loans

Variables	Transformation	Pearson	Variables	Transformation	Spearman	Variables	Transformation	Kendall
Debt Service Ratio	lag4	70.94%	Debt Service Ratio	lag1	64.60%	Debt Service Ratio	lag2	47.71%
House Price Index	yoy	-62.76%	Non-Farm Employment	qoq	-57.94%	Non-Farm Employment	yoy	-43.67%
New Home Units	yoy_lag4	-62.30%	Household Fin. Obligations	lag4	57.78%	Household Fin. Obligations	lag3	41.52%
Household Fin. Obligations	lag4	55.26%	Personal Savings	lag4	-50.50%	Personal Savings	lag4	-36.02%
Nominal GDP	yoy_lag1	-50.89%	New Home Sales	yoy_lag4	-49.41%	Unemployment Rate	yoy_lag1	35.49%

Panel C: Real Estate Loans

Variables	Transformation	Pearson	Variables	Transformation	Spearman	Variables	Transformation	Kendall
New Home Sales	ln	-75%	Unemployment Rate	raw	70.43%	Unemployment Rate	raw	52.52%
Commercial Real Estate Price	yoy_lag2	-69%	House Price Index	yoy	-63.46%	Unemployment Insurance Claims	raw	46.12%
House Price Index	yoy_lag2	-68%	UIC	raw	61.95%	House Price Index	yoy	-45.09%
Unemployment Rate	raw	66%	CRE Price Index	yoy_lag1	-58.33%	CRE Price Index	yoy_lag1	-41.85%
BBB Corporate Bond Yield	lag4	50%	New Home Units	raw	-53.75%	New Home Units	raw	-38.60%

Table 5 Feature Selection C&I Loans

Table 5 reports the variables selected for C&I loans by stepwise regression, lasso regression (Panel A), and tree-based models GBM and RF (Panel B), with their standardized coefficients and variable importance.

Panel A: Stepwise and Lasso

Stepwise			Lasso		
Variables	Transformation	Std. Coef.	Variables	Transformation	Std. Coef.
Real GDP Growth	yoy_lag1	8.827	CRE Price Index	yoy	1.362
Nominal GDP Growth	yoy_lag1	5.281	New Home Unit	yoy_lag4	0.520
Non-Farm Employment	yoy	1.678	House Fin. Obligations	lag4	0.108
CRE Price Index	yoy	1.200	Prime Rate	yoy	0.097
New Home Unit	yoy_lag4	0.483	BBB Corporate Yield	lag4	0.057

Panel B: GBM and RF

GBM			Random Forest		
Variables	Transformation	Importance	Variables	Transformation	Importance
Non-Farm Employment	yoy	0.719	Non-Farm Employment	yoy	0.712
BBB Corporate Yield	lag4	0.138	BBB Corporate Yield	lag4	0.131
New Home Unit	yoy_lag4	0.049	New Home Unit	yoy_lag4	0.033
Market Volatility Index	lag2	0.023	House Fin. Obligations	lag4	0.024
CRE Price Index	yoy	0.022	Real GDP Growth	yoy_lag1	0.022

Table 6 Feature Selection Consumer Loans

Table 6 reports the variables selected for consumer loans by stepwise regression, lasso regression (Panel A), and tree-based models GBM and RF (Panel B), with their standardized coefficients and variable importance.

Panel A: Stepwise and Lasso

Stepwise			Lasso		
Variables	Transformation	Std. Coef.	Variables	Transformation	Std. Coef.
Real GDP Growth	yoy_lag3	5.195	Debt Service Burden	lag4	0.334
CRE Price Index	yoy_lag2	2.588	UR	raw	0.100
Nominal GDP Growth	yoy_lag1	2.137	House Fin. Obligations	lag4	0.028
New Home Unit	yoy_lag4	1.231	House Price Index	yoy	0
DJ Index	yoy_lag3	0.611	New Home Unit	yoy_lag4	0

Panel B: GBM and RF

GBM			Random Forest		
Variables	Transformation	Importance	Variables	Transformation	Importance
House Price Index	yoy	0.601	House Price Index	yoy	0.520
Debt Service Burden	lag4	0.122	Debt Service Burden	lag4	0.140

DJ Index	yoy_lag3	0.096	UR	raw	0.093
Real GDP Growth	yoy_lag3	0.042	New Home Unit	yoy_lag4	0.054
UR	raw	0.036	CRE Price Index	yoy_lag2	0.049

Table 7 Feature Selection Real Estate Loans

Table 7 reports the variables selected for real estate loans by stepwise regression, lasso regression (Panel A), and tree-based models GBM and RF (Panel B), with their standardized coefficients and variable importance.

Panel A: Stepwise and Lasso

Stepwise			Lasso		
Variables	Transformation	Std. Coef.	Variables	Transformation	Std. Coef.
House Price Index	yoy_lag2	4.167	House Price Index	yoy_lag2	1.902
Real GDP Growth	yoy_lag4	2.499	CRE Price Index	yoy_lag2	1.643
Non-Farm Employment	log	2.077	New Home Unit	log	1.275
CRE Price Index	yoy_lag2	1.701	BBB Corporate Yield	lag4	0.121
New Home Unit	log	1.423	UR	raw	0.049

Panel B: GBM and RF

GBM			Random Forest		
Variables	Transformation	Importance	Variables	Transformation	Importance
UR	raw	0.374	CRE Price Index	yoy_lag2	0.233
New Home Unit	log	0.180	UR	raw	0.183
BBB Corporate Yield	lag4	0.145	New Home Unit	log	0.166
DJ Index	log	0.081	House Price Index	yoy_lag2	0.165
House Price Index	yoy_lag2	0.065	Real GDP Growth	yoy_lag4	0.083

VII. Model Selection and Prediction

Table 8 compares the predictive performance of four models—Stepwise Regression, Lasso Regression, Gradient Boosted Machine (GBM), and Random Forest (RF)—across three loan portfolios. Using quarterly data for training (1990 Q1–2021 Q3) and testing (2021 Q4–2024 Q3), tree-based models (GBM and RF) consistently outperform regression-based models (Table 8).

For real estate loans, GBM achieves the best results, with the lowest in-sample MSE (0.0003) and out-of-sample MSE (0.0068), demonstrating its strong generalization and ability to capture nonlinear relationships. RF follows closely (out-of-sample MSE: 0.0191), while regression models have higher errors due to limited capacity to model data complexity.

In consumer loans, RF slightly outperforms GBM (out-of-sample MSE: 0.3416 vs. 0.3449), confirming tree-based models' effectiveness in modeling charge-off rates. Lasso performs better than Stepwise regression but worse than tree-based models.

For C&I loans, GBM again leads with the lowest in-sample (0.0009) and out-of-sample (0.0091) MSE, followed by RF (out-of-sample MSE: 0.0265). Tree-based models significantly outperforms regression methods, highlighting their ability to capture complex interactions.

Overall, GBM and RF demonstrate superior predictive accuracy and robustness across all loan portfolios, proving their effectiveness in modeling credit losses during volatile economic periods.

Table 8 Model Performance Comparison for Different Loan Portfolio

Table 8 compares the performance of predictive models (Stepwise, Lasso, GBM, and RF) for three loan portfolios using quarterly data from 1990 Q1 to 2024 Q3. The in-sample period is from 1990 Q1 to 2021 Q3, while the out-of-sample period is from 2021 Q4 to 2024 Q3 (12 quarters). The table reports the in-sample and out-of-sample Mean Squared Errors (MSE) for each model.

Portfolio	Model	In-Sample MSE	Out-of-Sample MSE
Commercial Real Estate	Stepwise	0.0929	0.2843
Commercial Real Estate	Lasso	0.0999	0.2471
Commercial Real Estate	GBM	0.0003	0.0068
Commercial Real Estate	RF	0.009	0.0191
Consumer Loans	Stepwise	0.3009	0.7738
Consumer Loans	Lasso	0.7848	0.6341
Consumer Loans	GBM	0.0022	0.3449
Consumer Loans	RF	0.0173	0.3416
Commercial & Industrial Loans	Stepwise	0.0739	0.0342
Commercial & Industrial Loans	Lasso	0.0812	0.0718
Commercial & Industrial Loans	GBM	0.0009	0.0091
Commercial & Industrial Loans	RF	0.0063	0.0265

Using the best models identified in Table 8, we conduct in-sample testing (1990 Q1 to 2021 Q3) and out-of-sample testing (2021 Q4 to 2024 Q3) for each portfolio. The actual charge-off rates alongside the model predicted charge-off rates are presented in Figures 4, 5, and 6.

Figure 4 shows the GBM model's performance in predicting charge-off rates for the C&I loans. The GBM model demonstrates strong predictive accuracy and effectiveness in capturing trends and fluctuations. The close alignment between the model's predictions and actual data during both in-sample period and out-of-sample period highlights its robustness and reliability for forecasting.

Figure 5 illustrates the RF model's predicted charge off rate for the consumer loans, comparing them to actual charge-off rates. The model accurately reflects charge-off rate trends, effectively capturing the cyclical and volatile nature of consumer loan performance. Its strong generalization during the out-of-sample period demonstrates the RF model's ability to incorporate macroeconomic factors, such as unemployment rates and household financial obligations, to produce reliable forecasts even under varying economic conditions.

Figure 6 highlights the GBM model's performance in predicting charge-off rates for the real estate loans. The model exhibits high accuracy, closely mirroring actual charge-off rates and successfully capturing cyclical trends and volatility. Its effectiveness in incorporating complex macroeconomic indicators, such as house price indices and unemployment rates, further emphasizes its reliability for forecasting the performance of real estate loans under diverse economic scenarios.

Figure 4

Predictions for the C&I loans using the GBM model. The figure compares actual values, in-sample predictions, and out-of-sample predictions. The training period is from 1990 Q1 to 2021 Q3, and the testing period is from 2021 Q4 to 2024 Q3.

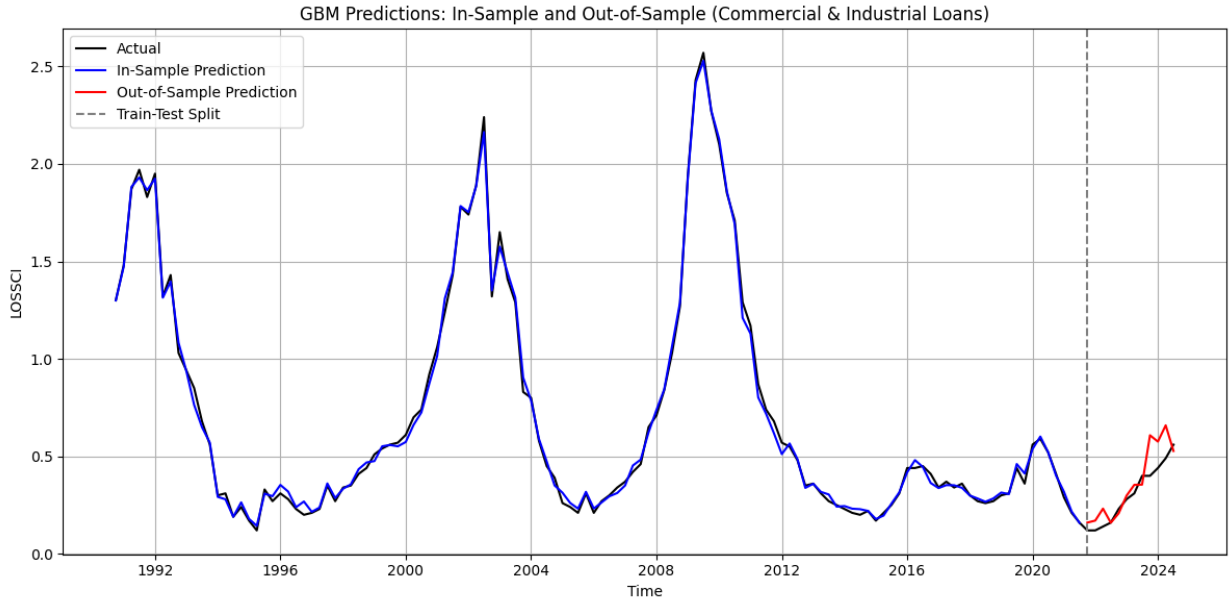


Figure 5

Predictions for the consumer loans using the RF model. The figure compares actual values, in-sample predictions, and out-of-sample predictions. The training period is from 1990 Q1 to 2021 Q3, and the testing period is from 2021 Q4 to 2024 Q3.

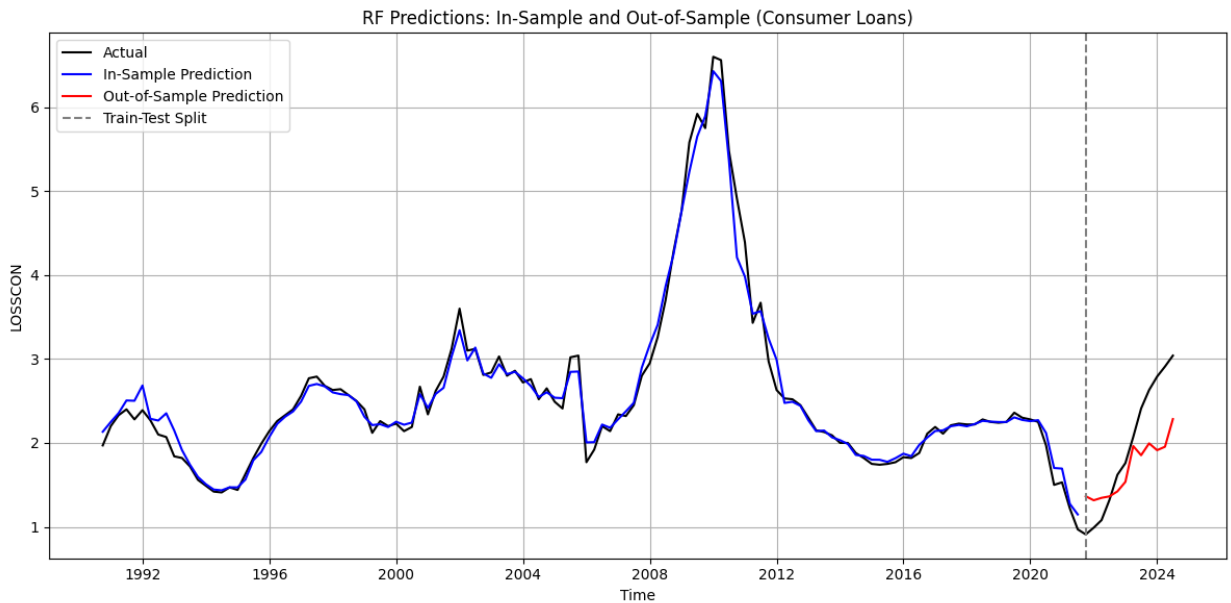
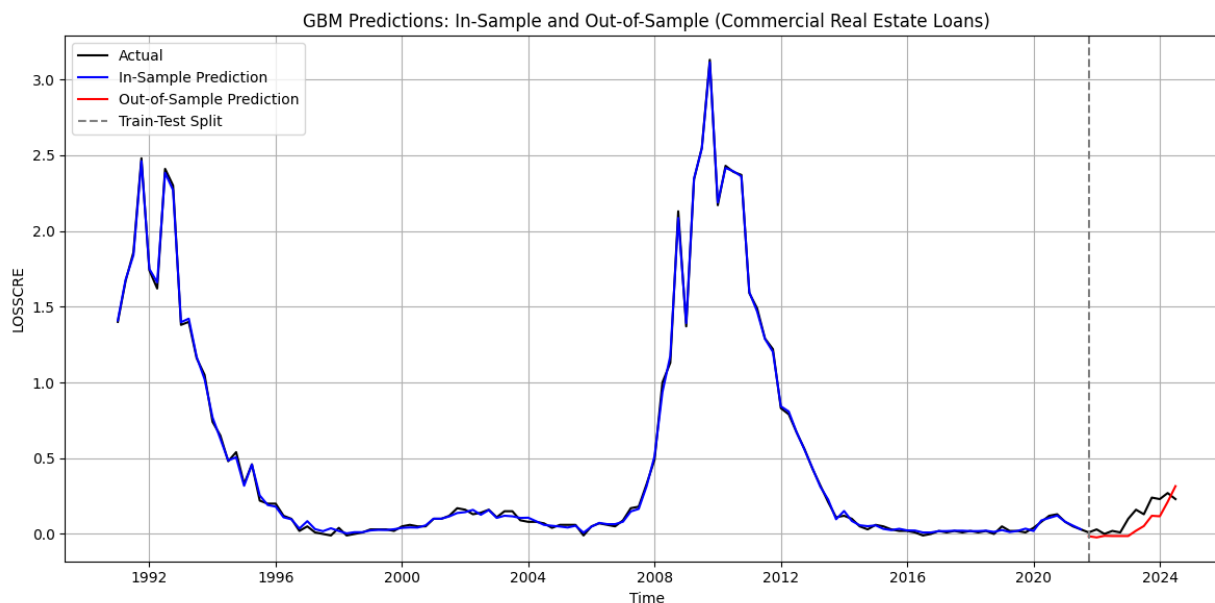


Figure 6

Predictions for the real estate loans using the GBM model. The figure compares actual values, in-sample predictions, and out-of-sample predictions. The training period is from 1990 Q1 to 2021 Q3, and the testing period is from 2021 Q4 to 2024 Q3.



VIII. Results Discussion and Implications to Credit Risk Modelling

Credit risk management becomes a critical focus for financial institutions, particularly following the 2008–2009 global financial crisis (GFC) and the unprecedented volatility introduced by the COVID-19 pandemic. This study examines the relationship between macroeconomic factors and charge-off rates for C&I loans, consumer loans, and real estate loans issued by U.S. commercial banks from 1990 to 2024. Our findings provide a comprehensive analysis of macroeconomic risk factors that can be effectively used in credit loss forecasting (LF) and CECL models.

The results highlight key macroeconomic variables that have the strongest predictive power for credit loss forecasting across different loan portfolios. For C&I loans, non-farm employment, CRE price index, and new housing units emerged as the most significant predictors, emphasizing the importance of labor market and commercial real estate dynamics. For consumer loans, house price index, debt service burden, and household financial obligations are the most critical variables, reflecting the influence of household financial conditions and housing market trends. Finally, for real estate loans, house price index, new housing units, and CRE price index demonstrate the strongest predictive power, underscoring the dominance of housing market variables in determining real estate loan performance.

By incorporating these macroeconomic factors, banks can improve the robustness of their credit risk models, particularly under stress-testing scenarios. This study also demonstrates the advantages of combining feature selection techniques with traditional correlation analyses to identify the most relevant macroeconomic variables. Importantly, the inclusion of COVID-19-era data provides a unique opportunity to assess the stability and relevance of these relationships during periods of extreme economic stress.

Given that most banks primarily rely on limited datasets covering a single business cycle (e.g., GFC), our long-term analysis spanning multiple business cycles—including the COVID-19 pandemic—offers a more robust understanding of the dynamic relationship between macroeconomic factors and credit loss. While our findings are robust, they should be interpreted with caution due to potential differences between the COVID-19 pandemic and prior recessions. Nonetheless, this research provides practical insights for banks, regulators, and policymakers to enhance credit risk modeling frameworks and to design more effective stress-testing and loss forecasting methodologies.

Reference

- Agarwal, S., & Liu, C. (2003). Determinants of credit card delinquency and bankruptcy: Macroeconomic factors. *Journal of Economics and Finance*, 27(1), 75-84.
- Ausubel, L. M. (1997). Credit card defaults, credit card profits, and bankruptcy. *American Bankruptcy Law Journal*, 71, 249.
- Bellotti, T., & Crook, J. (2013). Forecasting and stress testing credit card default using dynamic models. *International Journal of Forecasting*, 29(4), 563-574.
- Betz, J., Krüger, S., Kellner, R., & Rösch, D. (2020). Macroeconomic effects and frailties in the resolution of non-performing loans. *Journal of Banking & Finance*, 112, 105-212.
- Bonfim, D. (2009). Credit risk drivers: Evaluating the contribution of firm level information and of macroeconomic dynamics. *Journal of Banking & Finance*, 33(2), 281-299.
- Breeden, J. L., & Crook, J. (2020). Multihorizon discrete time survival models. *Journal of the Operational Research Society*, 73(1), 56-69.
- Castro, V. (2013). Macroeconomic determinants of the credit risk in the banking system: The case of the GIPSI. *Economic Modelling*, 31, 672-683.
- Djeundje, V. B., & Crook, J. (2018). Incorporating heterogeneity and macroeconomic variables into multi-state delinquency models for credit cards. *European Journal of Operational Research*, 271(2), 697-709.
- Figlewski, S., Frydman, H., & Liang, W. (2012). Modeling the effect of macroeconomic factors on corporate default and credit rating transitions. *International Review of Economics & Finance*, 21(1), 87-105.
- Fung, T., & Wong, M. (2002). Modeling credit card charge-off ratios: The case of Hong Kong. *City University of Hong Kong: Department of Economics & Finance*.
- Giesecke, K., Longstaff, F. A., Schaefer, S., & Strebulaev, I. (2011). Corporate bond default risk: A 150-year perspective. *Journal of Financial Economics*, 102(2), 233-250.
- Gross, D. B., & Souleles, N. S. (2002). Do liquidity constraints and interest rates matter for consumer behavior? Evidence from credit card data. *The Quarterly journal of economics*, 117(1), 149-185.
- Hackbarth, D., Miao, J., & Morellec, E. (2006). Capital structure, credit risk, and macroeconomic conditions. *Journal of Financial Economics*, 82(3), 519-550.
- Jakubik, P. (2007). Macroeconomic environment and credit risk. *Czech Journal of Economics and Finance (Finance a úvěr)*, 57(1-2), 60-78.
- Koju, L., Koju, R., & Wang, S. (2019). Macroeconomic determinants of credit risks: evidence from high-income countries. *European Journal of Management and Business Economics*. 29(1), 41-53.

- Liu, J., & Xu, X. E. (2003). The predictive power of economic indicators in consumer credit risk management. *Rma Journal*, 86(1), 40-45.
- Liu, Z., Pu, X., & Zhao, X. (2015). What Moves the Correlation between the Equity and Credit Default Swap Markets? *The Journal of Fixed Income*, 25(2), 72-87.
- Mählmann, T. (2005). Biases in estimating bank loan default probabilities. *The Journal of Risk*, 7(4), 75-102.
- Palese, P. (2004). The great influenza: The epic story of the deadliest plague in history. *The Journal of Clinical Investigation*, 114(2), 146-146.
- Pesaran, M. H., Schuermann, T., Treutler, B. J., & Weiner, S. M. (2006). Macroeconomic dynamics and credit risk: a global perspective. *Journal of Money, Credit and Banking*, 38(5), 1211-1261.
- Pu, X., & Zhao, X. (2012). Correlation in credit risk changes. *Journal of Banking & Finance*, 36(4), 1093-1106.
- Stavins, J. (2000). Credit card borrowing, delinquency, and personal bankruptcy. *New England Economic Review*, 15-30.
- Taghiyeh, S., Lengacher, D. C., & Handfield, R. B. (2021). Loss rate forecasting framework based on macroeconomic changes: Application to US credit card industry. *Expert Systems with Applications*, 165, 113954.
- Virolainen, K. (2004). Macro stress testing with a macroeconomic credit risk model for Finland. *Bank of Finland Research Discussion Paper*, 18/2004.

Appendix A: Variable Explanations and FRED Data Series Codes

Variable Name	Description	Source	FRED Data Series Code
Commercial Real Estate Charge-Off Rate	Quarterly percentage of commercial real estate loans charged off. Reflects credit risk in CRE markets.	Federal Reserve Board	CORCREXFA CBS
Consumer Loans Charge-Off Rate	Quarterly percentage of consumer loans (e.g., credit cards, personal loans) charged off.	Federal Reserve Board	CORCACBS
Commercial and Industrial Loans Charge-Off Rate	Quarterly percentage of commercial and industrial loans charged off. Reflects credit risk in business lending.	Federal Reserve Board	CORBLACBS
Real GDP Growth	Quarterly percentage change in real GDP, representing overall economic activity.	U.S. Bureau of Economic Analysis	GDPC1
Nominal GDP Growth	Quarterly percentage change in nominal GDP, measuring the overall value of goods and services.	U.S. Bureau of Economic Analysis	GDP
Real Disposable Income	Quarterly real disposable personal income, adjusted for inflation.	U.S. Bureau of Economic Analysis	DPIC96
Nominal Disposable Income	Quarterly nominal disposable personal income, not adjusted for inflation.	U.S. Bureau of Economic Analysis	DPI
Unemployment Rate (UR)	Percentage of unemployed individuals in the labor force.	U.S. Bureau of Labor Statistics	UNRATE
CPI Inflation	Consumer Price Index for all urban consumers, reflecting overall inflation trends.	U.S. Bureau of Labor Statistics	CPIAUCSL
3-Month Treasury Rate	Yield on 3-month U.S. Treasury bills, reflecting short-term interest rates.	Federal Reserve Board	TB3MS
5-Year Treasury Rate	Yield on 5-year U.S. Treasury notes, representing medium-term interest rates.	Federal Reserve Board	GS5
10-Year Treasury Rate	Yield on 10-year U.S. Treasury notes, representing long-term interest rates.	Federal Reserve Board	GS10
BBB Corporate Yield	Yield on BBB-rated corporate bonds, indicating credit market conditions.	ICE BofA	BAMLC0A4C BBBEY
Mortgage Rate	Average interest rate on 30-year fixed-rate mortgages in the U.S.	Federal Reserve Board	MORTGAGE 30US
Prime Rate	Interest rate charged by banks to their most creditworthy customers.	Federal Reserve Board	WPRIME
Dow Jones Index	Dow Jones Industrial Average, reflecting overall stock market performance.	Dow Jones & Company	DJIA
House Price Index	All-transactions house price index for the U.S., reflecting real estate market conditions.	Federal Housing Finance Agency	USSTHPI
CRE Price Index	Commercial real estate price index, capturing trends in CRE markets.	Federal Reserve Board	BOGZ1FL075 035503Q
Market Volatility Index (VIX)	A measure of expected market volatility derived from S&P 500 index options.	Chicago Board Options Exchange	VIXCLS
Non-Farm Employment	Number of employed individuals outside the agricultural sector, reflecting labor market strength.	U.S. Bureau of Labor Statistics	PAYEMS

Retail Sales - Motor	Retail sales in the automotive sector, an indicator of consumer demand.	U.S. Census Bureau/Federal Reserve Board	DAUTOSA
New Housing Units	Number of new privately-owned housing units authorized in permit-issuing places.	U.S. Census Bureau	PERMIT
Personal Consumption Expenditures (PCE)	Total expenditures by households on goods and services.	U.S. Bureau of Economic Analysis	PCE
Personal Saving Rate	Percentage of disposable personal income saved by households.	U.S. Bureau of Economic Analysis	PSAVERT
Spot Crude Oil Price (WTI)	Price of West Texas Intermediate crude oil per barrel.	U.S. Energy Information Administration	WTISPLC
Initial Claims	Weekly initial claims for unemployment benefits, capturing labor market stress.	U.S. Employment and Training Administration	ICSA
Producer Price Index (PPI)	Index measuring average change over time in the selling prices received by domestic producers for their output.	U.S. Bureau of Labor Statistics	PPIACO
Debt Service Burden	Ratio of debt payments to disposable income, indicating consumer financial health.	Federal Reserve Board	TDSP
Household Financial Obligations	Financial obligations as a percentage of disposable income, covering mortgages, rent, and debt payments.	Federal Reserve Board	FODSP

Corporate Governance and Cash Holdings Decisions by U.S. Life Insurance Companies

Raja Bouzouita*

Abstract

In the context of the principal-agent framework, this article examines the impact of corporate governance on U.S. life insurers' cash holdings. A dynamic panel model estimates the factors that affect cash holdings. In the second stage, regression models estimate and test the influence of corporate governance variables on the adjustment speed toward target cash holdings. It is found that proxies for corporate governance such as board size and member independence are significant factors explaining the cash holdings decision and are effective in mitigating the principal-agent conflict. Furthermore, the adjustment rate to target cash holdings is estimated. The analysis shows that the adjustment rate varies across U.S. life insurance companies with the corporate governance proxies used. Finally, this study confirms the importance of other control variables in prior insurance studies analyzing the cash holdings decision.

Keywords: Insurance Industry, Cash Holdings, Corporate Governance.

JEL Classification: C23, G22, G30

I Introduction

Cash holdings decisions are paramount in the financial services industry. The life insurance industry is no exception. Life insurance companies hold cash assets to pay claims such as death benefits, disability benefits, and annuity payments. An additional need for cash holdings is required for some life insurance products such as cash value policies that give the right to policyholders to borrow against their cash surrender value. The role of insurance companies in the economy is important for their stakeholders/policyholders who expect and rely on them to pay death benefits, income replacement, long-term care, and retirement income. This vital role of the insurance industry makes several stakeholders interested in the corporate governance of the industry and its decision cash holdings. Corporate governance refers to the systems and mechanisms that ensure effective control and management of a firm.

The complexity of the insurance business reduces the chance for all stakeholders (shareholders and policyholders) to monitor insurance companies' managers' decisions for several reasons. First, the insurance industry is heavily regulated. Often the stated goal of insurance regulation is solvency in addition to remedial to certain insurance market failures such as asymmetric information and market conduct (Klein, 2012). These market failures may lead insurance companies to take excessive risk further exacerbating the principal/agent conflict. Second, customers find insurance products very complex and cannot judge the financial risk their companies represent. Third, the insurance industry business is opaque because it is difficult for investors to judge and assess the quality of the assets and liabilities of insurance companies (Sommer et al., 2006; Chiang et al., 2022). Insurance companies' liabilities are complex and calculated as the present value of future claims that could extend decades into the future and

* University of Central Missouri, College of Business and Professional Studies, bouzouita@ucmo.edu

depend on several factors such as the mortality rate, cost of health care, and the expected interest rate.

Under perfect market conditions, companies will hold enough cash to meet their financial obligations until new expected sources of operating cash flows are made available. However, markets are far from being perfect. Cash holdings are retained to mitigate some market imperfections or adverse events. The early explanations of cash holdings decisions centered on the transactions, precautionary, and speculative motives (Keynes, 1936; Baumol, 1952). Subsequent theories address the cash-holding decisions in the context of corporate governance. The mainstream literature on corporate governance and cash holdings focuses on the principal-agent conflict of interest. Managers are inclined to hold disproportionate cash to pursue their own interests. The existence of excessive cash may exacerbate the agency-principal conflict. This paper analyzes corporate governance's impact on cash-holding decisions in U.S. life insurance companies. Understanding the effect of corporate governance will help enhance insurance companies' profitability by eliminating inefficiencies emanating from poor governance. A profitable life insurance industry benefits consumers and insurance firms' investors. Unlike industrial firms, liquidity management is essential to insurance companies that have an obligation to pay their policyholders' claims. This study will shed light on this decision in the context of agency theory while considering other variables that are found important in previous studies. The impact of corporate governance on the cash-holding decision in life insurance has not been studied thus far.

The remainder of the paper is organized as follows. Section 2 reviews prior studies on determinants of cash holdings in the insurance industry. Section 3 examines the theoretical background on the effect of corporate governance and cash holdings. Section 4 discusses the data, variables, and estimation procedure. Section 5 presents the empirical results. Section 6 provides a summary and concluding remarks.

II LITERATURE REVIEW

The literature on cash holdings is extensive. One of the cornerstones of financial management is securing adequate liquidity to meet day-to-day obligations. Only studies directly addressing cash-holding in the insurance industry will be reviewed in this section. The first study on this topic was by Wells, Cox, and Gaver (1995) who focused on the relationship between the life insurance corporate structure and the level of cash flow. The authors find that mutual companies have the propensity to hold higher levels of free cash than stock companies. Among other significant variables were size measured by total assets and leverage. One shortcoming of the study is its use of a single-year (1989) sample. Colquitt, Sommer, and Godwin (1999) study the other insurance industry segments using a multiple-year sample of property-liability insurers. Contrary to Wells, Cox, and Gaver (1995), the authors find that stock companies hold more cash than mutual companies. Similarly, the authors find that size, financial strength, leverage, and group affiliation are significant determinants of cash holdings. Both of the studies explain their advanced hypothesis within the principal-agent framework. Hsu, Huang, and Lai (2015) explicitly focus on corporate governance and cash in the property and liability insurance industry. They use corporate control proxies such as the proportion of outside directors, board size, percentage of shares held by directors, whether the board of directors is headed by the CEO, and other control variables consistent with Colquitt, Sommer, and Godwin (1999). The authors find evidence that the corporate governance variables contribute to alleviating the principle agent conflict. However, the

study includes only stock publicly traded property-liability insurance companies. As previous studies have shown, the corporate structure in insurance is characterized by both mutual and stock companies and this corporate structure is a source of conflict of interest between managers and other stakeholders. Second, a limited number of insurance companies are publicly traded. Having a sample that includes only publicly traded companies omits a large representation of the whole industry. The next study by Xie, Wang, Zhao, and Lu (2017) builds on Wells, Cox, and Gaver (1995) by using a more refined classification of corporate structure in the insurance industry. Using sample panel data of property-liability insurers from 1993-2011, the authors find that mutual companies hold more cash than stock companies. The last study by Che, Fier, and Liebenberg (2019) analyzes the cash holdings of property-liability insurance companies, emphasizing the effect of market structure on cash holdings. The authors relying on the predation risk theory, find that U.S. property-liability insurers that operate in more concentrated markets hold more cash reserves.

This study adds to the existing literature on cash holdings by focusing on corporate governance in one segment of the insurance industry, the life insurance sector, and using a sample that includes both mutual and stock companies.

III INSURANCE BOARDS AND HYPOTHESES DEVELOPMENT

The monitoring role of the board of directors

The literature on corporate governance indicates that a company board can effectively mitigate the conflict of interest between managers and stakeholders. The conflict of interest arises from the diverging interests among stakeholders. Managers may prefer to maintain a high level of cash holdings to exploit for their benefits such as enjoying additional perks and overinvesting in real assets. The free cash flow theory states that managers of firms with excess funds, more than the optimal level, either tend to hold it as an excess surplus that would remain idle or earn a low return when invested in near cash assets. Alternatively, those managers may overinvest in real assets. While maximizing managers' utility, these overinvestments reduce the firm value as explained by Jensen and Meckling (1976) and Myers and Rajan (1998). Therefore, to reduce the impact of agency problems mainly overinvesting, increasing the board of directors' effectiveness by monitoring self-serving managers may achieve the desired alignment of managers' and shareholders' interests and enhancement of corporate value. The effectiveness of the board of directors depends on its composition and size. Usually, members could be either insiders or outsiders. Insiders are either executives or major stakeholders. It is argued that independent boards with a majority of outside members are more effective in supervising the CEO's decisions and better-aligning management interests with shareholders (Coles et al., 2014; Duchin et al., 2010). Independent directors are more likely to be motivated to play their role more effectively to bolster their reputation as valuable advisors and increase the demand for their skills. Another attribute of the board composition is the relationship between the board and the chief executive officer (CEO) which is related to board independence. In some companies, the CEO is also the board's chairperson. Consolidating these functions will grant the CEO greater control over the decision-making process and undue influence in appointing board members who will be less likely to challenge the CEO's decisions (Booth et al., 2002) further exacerbating the principal-agent conflict.

The size of the board is another characteristic that may affect its effectiveness. A large board may be necessary to have individuals representing different areas of expertise who will make good decisions. However, a large board may be slow in making decisions due to a lack of consensus. Therefore, a small board could be more effective than a larger one Jensen (1993) demonstrates that having small boards makes them more effective.

Based on the literature on corporate governance and cash holdings the following hypotheses are proposed:

H₁: Other things being equal, the higher the proportion of independent directors the lower the cash holdings.

One attribute of the board of directors is their relationship with the company, which could influence the effectiveness of their monitoring role of executives' actions. Independent directors will push for firm value-enhancing projects and therefore lower the cash holdings. However, the board of directors who are officers/ executives may be more willing to align with the CEO's decision and assume a less effective out of fear of reprisal such as being passed for a promotion.

H₂: Other things being equal, the higher the number of directors the less effective the monitoring role, hence the higher the cash holdings.

The company board's size is commensurate with its size. Understandably, boards include a variety of individuals with expertise in different areas. Large boards may not necessarily be more effective than smaller boards. Corporate boards with few directors may result in individual directors being more likely to work as hard as other board members. While on large boards, some members may not exert their best effort (slack) and the board may suffer from a free-rider problem. In addition, large boards may make it harder to achieve consensus and the decision-making process may suffer. Therefore, small boards are expected to be more effective than larger boards in resolving the principal-agent conflict.

H₃: Other things being equal, companies whose CEO is also the board's chair will have higher cash holdings.

The role of the corporate boards is to exercise oversight of management including the CEO. In many cases, the CEO is also the board's chair. In this case, the CEO manages both the company and the board. The CEO sets the agenda and controls the flow of information given to the board. The dual role strengthens the discretionary power of the CEO. Therefore, a CEO, with significant power and influence, might choose to stockpile large amounts of company cash rather than reinvesting it for growth, potentially harming shareholder value, and making it challenging for the board of directors to effectively monitor and control such behavior.

Sample, Variables, and Econometric Model

Data

Insurance companies must file their financial statements with the state's insurance department where they are licensed to operate. The individual state data is compiled by the NAIC (National

Association of Insurance Commissioners). This study uses life insurance statutory financial statements compiled by the NAIC. The data covers the period from 2010-2013. The data is reported on an annual basis. The sample was screened for companies with outliers such as zero policyholders' surplus, and negative premiums. This screening process resulted in an unbalanced panel set with 3201 company-year observations.

Variables

Following previous studies, the dependent variable, CashHoldings, is defined by the ratio of cash and short-term assets divided by total invested assets. Only invested assets will generate future revenues for the company. The amount of cash and invested assets is deducted from total assets. The corporate governance variables, the focus of this study will be discussed first.

BoardSize is defined as the total number of directors reported in the "Directors and Officers" file of the NAIC Data. As discussed earlier, the role of the board of directors is to set company strategies and supervise managers' decisions. The board members are usually knowledgeable of the industry and have legal, financial, and actuarial skills which would result in large boards that would make it hard to effectively monitor managers' decisions. It is predicted that the larger the size of the board of directors the higher the cash holdings.

The board of directors' effectiveness in monitoring managers' decisions depends on its composition whether the board is insider-controlled or outsider-controlled. Insider board members are either officers of the company or are major stakeholders. For this study, due to data availability, an insider board member is also an officer of the company. The IndepDirectors variable is measured by the proportion of independent directors to the total number of directors. It is hypothesized that the more independent directors the lower the cash holdings.

The final variable to account for corporate governance is Duality. This variable takes on the value one when the CEO is also the chair of the board of directors and zero otherwise. This arrangement of CEO duality leadership has some advantages. In this case, the CEO/Chair has more latitude (autonomy) in the decision-making process and may diminish the board of directors' effectiveness in its monitoring role. Therefore, I expect a positive relationship between the cash holdings and the variable Duality.

Few insurance companies are publicly traded. Being a public company could add a layer of capital market monitoring in mitigating the agent-principal conflict of interest. These companies have to meet the requirements of the listing stock exchange and securities laws that are meant to strengthen corporate governance and better align the company's incentives with shareholders' goals. The variable Public is binary and takes the value one if the life insurance company is publicly traded or its parent company is publicly traded.

Similar to prior research covering the insurance industry (Wells, Cox, and Gaver, 1995; Colquitt, Sommer, and Godwin, 1999; Xie, Wang, Zhao, and Lu, 2017), the following control variables are included in this study.

Size: This variable is measured by the natural log of total assets that accounts for economies of scale. The hypothesis is that larger companies are better equipped to withstand any fluctuations in liquidity needs. Larger insurance companies tend to be more diversified hence, their risk exposure is lower than companies whose business is concentrated in a few lines of insurance. It is hypothesized that larger companies have lower cash holdings than smaller companies.

Risk: Risk refers to the volatility of operating cash flows. It is measured by the standard deviation of cash flows from operations over the previous 5 years. This variable is used to account for risk. Companies with higher variability in cash inflows and outflows which are less predictable will hold higher amounts of liquid assets. This variable explains the traditional precautionary motive for holding cash assets.

Dividends: This variable is defined as a binary variable that equals 1 if the life insurance company pays dividends to either shareholders or policyholders and zero if no dividends are paid. It is expected that the relationship between dividends and cash holdings is negative based on the premise that paying dividends is a signal of good governance by reducing the excess cash that managers might use inefficiently (Xie, Wang, Zhao, and Lu, 2017). Stockholders' dividends and policyholders' dividends are not guaranteed. Dividends policy is at the discretion of management and participating life insurance policies are not contractual but are illustrated in the prospectus offered to customers.

Profitability: This variable is measured as the ratio of net income to capital. Insurance companies with stable earnings will have less need for cash holdings. The profits will be used as a buffer against any volatility in cash disbursements. Profitable companies with steady earnings will enable them to build cash reserves to be relied upon to meet corporate objectives. Therefore, we expect a negative relationship between cash holdings and profitability.

Organization: The insurance industry is dominated by two organizational structures: mutual and stock companies. Mutual companies are technically owned by their policyholders. The declared goal of mutual companies is to provide insurance coverage to members at a low cost. While market discipline may necessitate profit-maximizing behavior on the part of stock companies, it is not evident that mutually-owned insurers and stock insurers pursue the same objectives. Cummins and Zi (1999) analyze ownership control issues in the insurance sector and find evidence that mutual companies have higher costs than stock companies due to a lack of effective mechanisms to control and discipline managers. Their study gives credence to the expense preference hypothesis, that managers may have objectives other than profitability in mutual companies. To control for the effect of ownership structure the relationship by including a dummy variable, Organization, takes the value of one if the insurer is a mutual and zero if the insurer is a stock company. It is hypothesized that mutual companies will hold more cash than stock companies.

Affiliation: Insurance companies are established as either single non-affiliated companies, or they can be a member of a group of companies under common management. Groups tend to be larger and more diversified across lines of business than single companies and have more resources and access to funds. Therefore, companies affiliated with a group can share expertise with fellow members on complex insurance problems and may get financial help if needed. It is conjectured that companies affiliated with an insurance group will hold less cash than single companies. Niehaus (2018) argues that within a life insurance group, a transfer of capital will be made to struggling member insurers to meet claim payments as a signal to rating agencies and policyholders.

Age: Defined as the difference between the year of observation and the business start date. The effect of age is related to the experience of the company. Newly formed companies seeking growth may have a higher need for cash than older companies that perhaps have reached a growth plateau.

Leverage: An additional risk measure defined by the ratio of net premium written to policyholders' surplus. High leverage will require higher amounts of cash holdings. Insurance companies with high leverage have a high volume of business and therefore will have to expect to pay higher claims. Generally, the degree of risk-taking is influenced by the attitude of insurance companies' owners and managers toward risk. Carson and Hoyt (1995) show that the probability of insolvency increases with leverage.

The life insurance companies' decision on how much liquid assets to hold depends on the nature of the assets held such as the ease with which these assets could be converted into cash to meet the liquidity needs. Life insurance companies hold a variety of assets with varying degrees of liquidity. Among these assets is the proportion of fixed assets, non-invested (NInvAsset), deemed the least liquid. Therefore, the higher the proportion of non-invested assets the higher the demand for cash and near-cash assets. On the other hand, another class of liquid assets such as common stocks (C.Stock) may be relied upon when cash needs turn out higher than expected. In this case, the higher the proportion of invested assets the lower the demand for cash holdings.

Econometric Model

Panel data analysis is employed since the sample is a combination of time series and cross-sectional data. Prior literature suggests that corporate board size is endogenous because board size might be jointly determined with firms' specific variables (e.g., Raheja (2005) and Linck, Jeffrey Netter, and Yang (2008)) Some variables such as the board size or its composition might be determined simultaneously with the dependent variable therefore, a System-GMM is utilized. The basic OLS method will yield biased and inconsistent estimates because of the presence of endogeneity in one of the independent variables (board size) shocks to the dependent variable (cash holdings).

The system-GMM is implemented using the Arellano and Bond estimator (1991). The method tests for autocorrelation has a null hypothesis of no autocorrelation and is applied to the differenced residuals. The sample includes all life insurance companies except those with extreme outliers such as negative premiums or zero policyholders' surplus. To avoid survivorship bias, the sample is not restricted to companies with complete data for all the years. Some companies drop out of the sample due to either merger, bankruptcy, or surrender of license and new companies join the industry. In addition, there is a low risk of selection bias as the sample includes all U.S. life insurance companies that are in operation or active.

IV ESTIMATION RESULTS

GMM Regression Model

Table 1 shows the statistics for these variables. On average, life insurance companies' cash holdings are about 16% of total assets which is lower than that of property liability insurers which is about 20% in 2011 (Xie, Wang, Zhao, Lu (2017)). Life insurers try to minimize their holdings of cash because most of their liabilities are long-term which makes the opportunity cost of cash

holdings very high. Life insurance claims are more predictable than those of property liability companies. The average size of the board is 6.77 and the number of independent directors is 4 which indicates that, on average more than half of the directors are not affiliated with the company they serve. The other corporate governance variable is the duality of CEO and chair positions. This explanatory variable shows that about a third of CEOs of life insurance companies also serve as the board chairperson. Several insurance companies are family-owned and managed by family members who serve as CEO and on the company board.

Table 1: Descriptive Statistics

Variable	Obs	Mean	Std.Dev	Min	Max
Cashholdings	3,123	0.16158	0.2449	0.0025	1
BoardSize	3,123	6.7951	3.2995	1	25
Ind.Directors	3,123	4.0935	3.5989	0	24
Duality	3,123	0.3129	0.4637	0	1
Public	3,123	0.2923	0.4656	0	1
Size	3,123	19.0627	2.9854	11.2797	26.6448
Risk	3,123	0.0016	0.0039	0	0.0005
Dividends	3,123	0.3153	0.4656	0	1
Profitability	3,123	0.1365	0.2034	-0.2686	0.9825
Organization	3,123	0.0563	0.2306	0	1
Affiliation	3,123	0.2305	0.4212	0	1
Age	3,123	50.2424	53.3184	1	171
Leverage	3,123	0.6563	0.3349	0.0177	2.061
NInvAssets	3,123	0.1343	0.201	0.2686	0.9825
C.Stock	3,123	0.0509	0.1312	0	0.9905

The estimation results of the system GMM estimator are reported in Table 2. For each regression, I indicate whether the coefficients are statistically different from zero (T-statistics); the first- and second-order correlation tests (AR1 and AR2); and the Hansen/Sargan tests of instrument validity. The statistical tests do not reject the validity of the model and confirm both the absence of second-order serial correlation and the validity of the instruments we use to avoid the endogeneity problem. The board of directors is the governing body of corporations, its role is to monitor management and protect stakeholders' interests. The analysis shows strong evidence in favor of corporate governance variables influencing the cash holding decision in the life insurance industry. In particular, the size of the board and the number of independent directors have a statistical and significant effect on the cash holdings of life insurance companies. The coefficient of the board size is positive and statistically significant. Companies with large boards hold higher levels of cash and near-cash assets than companies with smaller boards. This result indicates that large boards are less effective in mitigating the agency problem. A board with a large number of directors can struggle with coordinating actions, and effectively communicating information, and may potentially allow the CEO to exert too much influence due to the difficulty in managing such a large group, leading to potential issues with oversight and accountability. Similar to Xie, Wang, Zhao, and Lu (2017), this analysis finds that the board size variable shows that life insurance

companies with larger boards hold higher amounts of cash indicating that larger boards are ineffective in affecting managers' decisions.

Table 2: Corporate Governance Impact on Cash Holdings

Dependent	Model 1 GMM		Model 2 GMM	
	CashHoldings		CashHoldings	
Explanatory Variables	Coefficient	T-Stat	Coefficient	T-Stat
Board Size	0.0055***	2.59	0.0029**	2.16
Indep. Directors	-0.0080***	-3.25	-	-
Duality	-0.0061	-0.72	-0.0491	-0.57
Majority Indep.	-	-	-0.1549	-1.22
Public	0.0362**	2.89		
Size	-0.0408***	-6.66	-0.0408***	-6.68
Risk	0.0001***	5.77	-0.0002***	5.69
Dividends	-0.0156	-1.58	-0.0161	-1.63
Profitability	-0.0045***	-5.01	-0.0045***	-5.00
Organization	0.0467**	1.33	0.0553	1.58
Affiliation	0.0434**	2.40	0.0441**	2.44
Age	-0.0003***	-6.97	-0.0003***	-6.84
Leverage	-0.1101	-1.68	-0.1090	-1.58
NInvAssets	0.1642***	5.00	0.1639***	4.99
C.Stock	-0.2173***	-5.34	-0.2198***	-5.42
Constant	1.0252***	12.00	1.0123***	11.92
AR(1)	-3.85***		-3.83***	
AR(2)	0.42		0.43	
Sargan Test	6.69***		6.82***	

The table reports the parameter estimates of the GMM model, t-statistics are based on robust estimates of standard errors. AR(1) and AR(2) are tests of first and second serial correlation respectively. The null hypothesis no serial correlation in the first order is rejected. A first-order correlation is expected due to the presence of the lagged dependent variable in the model. The null hypothesis of no serial correlation in the second order is not rejected. The Sargan test of overriding restrictions is statistically significant.

*10% significant, **5% significant, ***1% significant.

The coefficient of the proportion of independent directors is negative and significant. The presence of independent directors the lower the cash holdings. This result supports the argument that having outside directors who by definition are not affiliated with the company, improves

management oversight and reduces the conflict of interest among stakeholders as predicted by theory. The results confirm the findings of Xie, Wang, Zhao, and Lu (2017) who find a similar result for publicly traded property and liability insurance companies. The results show that a company whose CEO is at the same time the chairperson of the board of directors has no bearing on the cash holdings. Other control variables such as size, risk measures, leverage, and dividend payments have the expected sign and are noticeably consistent with previous insurance studies such as Wells, Cox, and Gaver (1995) and Lee, Sommer, and Godwin (1999).

The results from the estimation of Model 2 use the proportion of independent directors if it exceeds 50 percent, under the assumption that the board of directors would be more effective if it had a majority of independent directors. The results show that this variable is not significant. Therefore, in the life insurance industry, a board of directors with a majority of independent directors does not affect cash holdings. This can be attributed to the following. Insurance companies' boards include several actuaries whose expertise is purely technical and therefore will be given more autonomy to devise and implement strategies that maximize shareholders' wealth. In conclusion, for life insurance companies a majority of outside directors is not needed to have an effective board to be successful in supervising management.

Adjustment Speed Estimates

Prior literature shows that companies hold target cash holdings based on the costs and benefits of maintaining cash ratios (Baumol, 1952; Tobin, 1956). Due to market imperfections, companies may deviate from the optimum level of cash holdings. Theoretically, companies have a target cash holdings level that minimizes costs and reflects management risk tolerance. For insurance companies having adequate liquidity levels is paramount (necessary) to honor policyholders' claims. Management's goal is to reach that optimum level of cash holdings which would allow the company to operate more efficiently and improve corporate performance. Firms with better corporate governance can adjust their cash holdings and reach their target levels. Excess cash holdings are wasteful to life insurers because most of their liabilities are long-term with predictable payouts. The estimated level of cash holdings in the previous section is used to calculate the adjustment speed for cash holdings. This section compares firms' cash holdings adjustments across corporate governance measures. The following model is estimated.

$$CashHolding_{it} - CashHoldings_{i,t-1} = \lambda(Cashholdings^*_{it} - CashHoldings_{i,t-1}) \quad (1)$$

where, $CashHoldings^*$ is the predicted level of cash (or target) and λ is the speed of adjustment.

In equation (1), the adjustment speed λ explains the actual changes in cash holdings from period t-1 to t, and it (λ) lies between 0 and 1, indicating multiple periods of adjustment. That is, the coefficient estimate for the lagged cash holdings represents the proportion of deviation closed in one period. Given equation (1), $\lambda = 1$ suggests full adjustment within one year for any deviation from target cash holdings. λ may be zero or close to zero in an extreme case of inertia or managerial indifference to cash holdings shocks. Generally, λ is expected to be more than zero but less than one. Whether life insurance companies can quickly adjust is an important research question because of the consequences of deviating from optimal levels of liquid assets, especially for financially constrained firms. Insurance companies' needs for cash and near-cash assets are paramount in fulfilling their promises of paying claims.

Table 3 presents the initial set of simple cash adjustment speed regressions following Equation (1), I estimate the cash ratio changes in a year against the deviation from the target cash ratio at the beginning of the year. The target cash ratio is the predicted value from the yearly cash ratio regression discussed above. The coefficient on the deviation from the target cash ratio is 0.2079, meaning that the estimated adjustment speed for the cash ratio is 20.79% each year. Firms close a substantial portion of the gap between actual and target cash ratios each year, indicating that managers actively consider target ratios. Yet the adjustment is far from complete (in which case the estimated adjustment speed would be 100%), likely because of the costs of making adjustments and possible agency problems.

Next, the asymmetry in adjustment speeds is estimated by splitting the sample into high cash excess holdings (75th percentile) and low excess cash holding (25th percentile). The second model of Table 3 presents the results from the split sample. The coefficient on the deviation from target cash is 0.2613 for the companies in the 25th percentile with a p-value of less than 0.001. The coefficient on the deviation from the target is 0.1767 for the companies in the 75th percentile. The higher adjustment speed at high excess cash levels indicates that companies with excess cash levels aspire to alleviate agency problems.

Table 3: Adjustment Speed of Cash Holdings

Independent Variable	Full Sample			Subsample with Excess Cash \leq 25%			Subsample with Excess Cash \geq 75%		
	β	T-stat	R ² (%)	β	T-stat	R ² (%)	β	T-stat	R ² (%)
Change in Cash	0.2079***	17.43	11.60	0.2613***	12.41	21.16	0.1767***	7.72	9.46
Constant	0.0002	0.11		-0.0139***	-8.13		-0.0082	-1.16	

Corporate Governance Effects on Cash Holdings Adjustments

Next, the asymmetry in adjustment speeds using regressions is augmented with an interaction variable between the deviation from the target ratio and the different corporate governance variables. Table 4 presents the results from this augmented regression. The first column of Table 4 reports the board size impact on the adjustment speed toward the target ratio of cash holdings. The coefficient on the deviation from target cash is 0.4233 with a p-value less than 0.001. The coefficient of the interaction variables board size \times deviation is negative and significant at the 1% level. The results suggest that companies with larger boards tend to adjust their cash holdings more slowly than those with smaller boards. This suggests that larger boards may be ineffective and make it harder to achieve consensus.

The second column of Table 4 shows the independent directors' adjustment speed estimates. The interaction term coefficient is positive and significant. The results suggest that companies whose board of directors is independent cash holdings adjust faster than companies having a board of directors dominated by insiders. The coefficient on the interaction term (dual \times deviation) is significantly negative, suggesting that companies whose CEO is also the chairperson of the board adjust more slowly toward the target cash level than companies where the CEO and board president positions are separate.

Overall, the results reported in Table 4 provide further evidence of corporate governance influence on cash holdings and whether these variables affect how life insurance companies adjust their cash holdings to the optimal level.

Table 4: Adjustment Speed and Corporate Governance Measures

Independent Variable	Board Size		Independent Directors		Duality	
	β	T-stat	β	T-stat	β	T-stat
Excess Cash	0.4233***	0.76	0.3468***	10.48	0.3639***	12.54
Interaction	0.0231***	4.41	0.0203***	4.16	-0.1071***	3.01
Constant	0.0037**	2.17	-0.0023	-1.48	-0.0027	-1.39
R ²	13.82%		13.66%		12.10%	

V Conclusion

Sound risk management dictates that life insurance companies maintain adequate liquidity to pay policyholders' claims promptly. In corporate governance, life insurance company boards play a role in mitigating the conflict of interest between different stakeholders. However, the structure and makeup of the insurance board, including its size, the independence of its members, and how effectively it operates, can directly influence how well it can fulfill its supervision role by actively monitoring management and providing strategic advice, which is crucial for ensuring sound financial practices and mitigating risk within the insurance company.

This paper investigates the relationship between life insurance companies' cash holdings and measures of corporate governance. The analysis provides strong empirical evidence for corporate governance's role in mitigating the conflict of interest between all stakeholders. The system GMM estimation shows a negative relation between the proportion of independent directors and cash holdings. This result supports the argument that adding outside directors to the board improves management monitoring and reduces the conflict of interest among stakeholders, as predicted by theory. The size of the board of directors has a positive and significant effect on the cash holdings. Larger boards are less effective in exercising their monitoring role due to the complexity of U.S. life insurance companies' operations. Firms with better corporate governance can adjust their cash holdings and reach their target levels. The second stage of the study estimates the adjustment speed toward the optimum level of cash holdings. The results show interesting interrelations between the adjustment speed and corporate governance variables used in this study. Most importantly, the adjustment speed is higher for companies whose board of directors includes more independent directors lending more credence to the importance of independent directors on life insurance boards. The research findings demonstrate a strong connection between robust corporate governance practices and the ability to effectively manage potential conflicts of interest arising from large cash reserves within financial services companies, suggesting that good governance mechanisms can help prevent managers from misusing excess cash at the expense of stakeholders. This research could inform regulatory bodies on how to design governance frameworks that incentivize effective board composition and functioning and its impact on liquidity management in the life insurance industry.

REFERENCES

- Arellano, M., & Bond, S. (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies*, 58, 277–297.
- Baumol, W. J. (1952). The transactions demand for cash: an inventory theoretic approach. *The Quarterly Journal of Economics*, 66, 545-5
- Che, X., Fier, S. G., & Liebenberg, A. P. (2019). The effect of predation risk on cash holdings: Empirical evidence from the US property-liability insurance industry. *Risk Management and Insurance Review*, 22(3), 329-358.
- Chiang, Chia-Chun; Kim, Hugh Hoikwang; Niehaus, Greg (2022) “Opaque Liabilities, Learning, and the Cost of Equity Capital for Insurers” *The Journal of Risk and Insurance*, Vol. 89 Issue 4, p1031-1076.
- Coles, J. L., Daniel, N. D., and Naveen, L. (2014). Co-opted boards. *The Review of Financial Studies*, 27(6), 1751-1796.
- Colquitt L. Lee, David W. Sommer, Norman H. Godwin (1999) “Determinants of Cash Holdings by Property-Liability Insurers” *The Journal of Risk and Insurance*, Vol. 66, No. 3, 401-415.
- Duchin, R., Matsusaka, J. G., and Ozbas, O. (2010). When are outside directors effective? *Journal of Financial Economics*, 96(2), 195-214.
- Hsu WenYen, Yenyu Huang, and Gene Lai (2015) “Corporate Governance and Cash Holdings: Evidence from the U.S. Property-Liability Insurance Industry” *The Journal of Risk and Insurance*, Vol. 82, No. 3, 715-748.
- Jensen, M.C. (1986), Agency costs and free cash flow, corporate finance and takeovers, *American Economic Review*, Vol. 76 No. 2, pp. 323-329.
- Jensen, M.C. (1993), The modern industrial revolution, exit, and the failure of internal control systems. *The Journal of Finance*, 48, 831-880.
- Jensen, M.C. and Meckling, W.H. (1976), “Theory of the firm: managerial behavior, agency costs, and ownership structure”, *Journal of Finance Economics*, Vol. 3 No. 4, pp. 305-360.
- Keynes J.M. (1936). *The General Theory of Employment, Interest, and Money*.
- Klein R. W (2012) “Principles for Insurance Regulation: An Evaluation of Current Practices and Potential Reforms” *The Geneva Papers*, 37, 1 75–199.
- James S. Linck, Jeffrey M. Netter, Tina Yang (2008) “The Determinants of Board Structure” *Journal of Financial Economics*, 78(2), 308-328.
- Niehaus G. (2018). “Managing Capital Via Internal Capital Market Transactions: The Case of Life Insurers” *Journal of Risk and Insurance*, Vol. 85, No1, 69-106
- Raheja G. Charn (2005), “Determinants of Board Size and Composition: A Theory of Corporate Boards” *Journal of Financial Quantitative Analysis*, 40(2), 283-306.
- Wells B. P. Larry A. Cox and Kenneth M. Gaver (1995) “Free Cash Flow in the Life Insurance Industry” *The Journal of Risk and Insurance*, Vol. 62, No1 pp. 50-66.
- Xie, Wang, Zhao, Lu (2017) Cash Holdings Between Private and Public Insurers – A Partial Adjustment Approach *Journal of Banking & Finance* Volume 82, Pages 80-97

Wherefore Art Thou?
Small business, healthcare coverage, and group association
 Roger Lee Mendoza*

Abstract

Small business/group research in health insurance is academically compelling given the dearth of literature, and insightful from both a theoretical and practical standpoint. The small group is a statutorily defined, government-regulated, and vital insurance market. We investigate why (and how) small firms manage to offer employee healthcare coverage despite the substantial costs and their employer mandate exemption. A corollary question is how they might band together and collaborate to further contain rising premium costs, gain leverage with insurers, and overcome mandatory protections and standards for small group plans. We find strategic differences between plans offered by small and large groups through which small groups might mitigate costs and trade-offs arising from employer propensities, plan design and benefits, financing, risk management, and associational options. Minimizing transaction costs of search and information, bargaining and decision-making, and monitoring and compliance, and market signaling within and between small firms to contain information asymmetries, are also key to achieving cost-efficiency in employer healthcare coverage.

Keywords: small business/group; employer-sponsored health insurance; Affordable Care Act (ACA); small group market; association health plan; actuarial risk

JEL Classification: G22, I11, I13, I18

I. INTRODUCTION

Employer-sponsored health insurance, alternatively called “group health insurance,” is the largest form of healthcare coverage in the United States, insuring 157 million people or over 58 per cent of the non-elderly population under age 65 besides their dependents (KFF, 2019). In a fully-insured health plan, the employer pays premiums to the insurer in exchange for the latter’s coverage of the costs of, and financial risks associated with, their employees’ healthcare over a defined period of time. The fully-insured group market includes both small group and large group health plans. It excludes employer-sponsored plans that are completely or partially self-funded (Morrisey, 2020).

The small group is a statutorily defined, government-regulated, and vital market in health insurance, with over 13 million covered lives (AHIP, 2019). In most states, a small group health plan is a group insurance product covering small business firms, not-for-profit organizations, and associations of any kind (e.g., local chamber of commerce or florists

*College of Business and Economics, California State University-Los Angeles, Los Angeles, CA 90032, Email: rmendoza@calstatela.edu

association) having one to 49 full-time equivalent employees (FTEs). In four states (California, Colorado, New York, and Vermont), a small group is capped at 100 FTEs. In any case, a small group with less than 50 FTEs is automatically exempt from the so-called “pay or play” mandate of the ACA (Affordable Care Act). Almost 97 percent of all U.S. firms have less than 50 FTEs and thus qualify for this employer mandate exemption (U.S. Census Bureau, 2016). The ACA’s employer mandate requires that any firm with 50 or more FTEs either provide “affordable” and “minimum value” healthcare to their workers and dependents, or else pay a hefty tax penalty toward a publicly-provided system that covers uninsured Americans (IRC, section 4980H, 2015).

The small group market remains under-explored and understudied. And of the five major health insurance markets in the U.S. (individual/non-group, small group, large group, Medicare Advantage, and Medicaid managed care), there has been little policy discussion about the fully-insured, small group market (Hall and McCue, 2018). This is particularly telling, considering that American business is overwhelmingly small group. Small groups with less than 10 workers make up 62 percent of all U.S. employers, while those with less than 20 workers represent over 89 percent of all U.S. employers (U.S. Census Bureau, 2016).

The scarce academic literature suggests that small group objectives and incentives behind healthcare coverage are employee attraction, retention, and productivity (Chamberlain and Tian, 2016), but often in a way that might also establish some comparative advantages to the organization and/or its owner. Cost — and its long-run risk implications — remains the prime disincentive to small groups (Long, Rae & Claxton, 2016; Mayer, 2020). Smaller risk pools and limited actuarial experience resulting in higher premiums, higher administrative costs per employee, and lack of dedicated staff to administer health plans make small groups much less likely to offer health insurance than larger employers (AHRQ, 2020). Yet, over half of small firms with less than 50 FTEs still offer health insurance (Miller, 2016; Hall & McCue, 2018).

Little else is known about small group health insurance beyond these incentives and challenges. In exploring organizational behavior, particularly information asymmetry, healthcare decision-making, and signaling between small businesses, the question for analytical investigation is why (and how) small firms still manage to offer healthcare coverage to their workers despite its substantial cost. A corollary question is how small firms might band together and collaborate to further contain cost and overcome mandatory standards and protections imposed by law on small group health plans. We seek to fill in the gaps in the insurance and risk management literature on small groups by illuminating the strategic choices, alternatives, and consequences of offering healthcare coverage either as a small business or through an association of small businesses. In doing so, we may further contribute to the literature on AHPs by distinguishing small groups from large group AHPs and clarifying their strategic advantages and limitations.

II. METHODS

The small firm or group (≤ 49 FTEs or ≤ 100 FTEs in the four previously named states) is our

unit of analysis, rather than their insured employees. Conversely, a large business or group has at least 50 or 101 FTEs, as the case may be. In exploring and analyzing organizational choices, decision-making, and communication, we compare small firms to large firms along several health insurance dimensions. The terms “business”, “firm”, and “group” are used interchangeably in this study. After all, these terms encompass any organization engaged in commercial, industrial, and/or professional activities in both private and public sectors, whether for-profit or not-for-profit entities.

We utilized two datasets to address our research questions. Our primary source was the SPSS-formatted, public-use dataset of the *Employer Health Benefits, Annual Survey* (EHB-AS) for five consecutive years (2015-2019) following full implementation of the ACA’s market reforms. This time frame also spans the period before and after the federal tax penalty for not carrying minimum health insurance, known as the individual mandate, was eliminated effective January 1, 2019 through the Tax Cuts and Jobs Act of 2017. The Kaiser Family Foundation and Health Research and Educational Trust (KFF/HRET) conducted during the five-year period 1,900 to 2,160 interviews annually with randomly selected non-federal public and private firms to generate single, point-in-time snapshots of employer-sponsored health benefits, plan costs, and selected health insurance issues. We disaggregated survey data, first by firm size so that our small firm sample conforms to the ACA, and then by particular insurance attributes. Small firm data was statistically tested for significance ($p < .05$).

Our second source was the one-year (2019) public-use database of AHPs generated by Association Health Plans, Inc., supplemented by interviews with some of their officers. Their survey corresponds to the last year of the EHB-AS dataset and offers a snapshot of the demographic characteristics of AHPs in 37 states that allowed them to fully operate. We similarly disaggregated survey data by association demographics and plan attributes.

Pertinent data was tabulated for analysis in the seven succeeding sections of this study. Owing to journal space restrictions, some tabulations had to be compressed and/or combined into one.

III. FINDINGS

Statutory incentives and constraints

The ACA applies essentially the same set of regulations to the small group market as it does the individual or non-group market from which the self-employed, uninsured, underinsured, and government-subsidized (but Medicaid-ineligible) low-income workers can buy healthcare coverage (Hall and McCue, 2018). Mandatory standards and protections for these plans include a minimum set of “essential health benefits”; the prohibition against using preexisting health conditions for enrollment eligibility and pricing; dependent children’s coverage until age 26; modified community rating (i.e., insurers cannot vary rates based on health status); a maximum pre-enrollment wait time of 90 days; guaranteed issue and renewal of insurance regardless of health status; and a limit on the percentage of premiums that can go towards insurer profit and overhead (known as the minimum “medical loss ratio”).

The ACA further established tax incentives for small firms to offer health insurance. Major drawbacks stem from, or in relation to, these statutory incentives. Healthcare tax credits equal to 50 percent of employer premiums are offered by the ACA to any sponsoring but small firm with less than 25 FTEs, which represents about 92.5 percent of all U.S. firms (U.S. Census Bureau, 2016). Tax credits are meant to encourage employer premium contributions and allow qualified small firms to better compete with large firms for skill and talent. However, the stipulated criteria is onerous and confusing to many small firms (Miller, 2016). By only offering to their FTEs plans under the ACA-created Small Business Health Options Program (SHOP), the mandatory average annual employee salary within a small group must be less than an annually indexed amount (\leq \$56,000 in 2025). In addition, the employer has to pay 50 percent or more of the premium for at least a single (employee-only) coverage plan. Tax credits are not applicable to the owner's premiums, exclude administrative costs beyond a certain yearly amount, and applies for a maximum of only two tax years. Hiring the 26th employee under this incentive scheme therefore has an immediate budgetary impact that disqualifies the employer from the tax credits. In terms of organizational behavior, the result is that these statutory credits are largely unused by small groups (Miller, 2016).

One hundred percent of employer-paid premium cost in any health plan (including health savings accounts and health reimbursement arrangements) is deductible from federal business taxes by an employer of any size. However, less than 20 percent of small business owners considers this as an incentive to offer health insurance. This is the case even among small firms in the four states with more (i.e., 50 to 100) workers who are subject to the employer mandate. The fast rising cost of healthcare coverage, including healthcare inflation, to the employer since the ACA's enactment could thus outweigh the tax benefit, although it may be an incentive to some employers (Miller, 2016).

Following the ACA's full implementation in 2014, small groups that have no more than 100 employees could obtain small group insurance from any of the following sources: 1) the federal SHOP exchange/marketplace; 2) by direct purchase from an insurance carrier/insurer; and 3) through an insurance broker or agent. Third party administrators may be hired by small groups to offer plan-related administrative services.

Before the ACA's passage, insurers routinely charged small firms higher premiums of 18 percent or more because it also costs more to insure smaller worker pools, whose collective risk tends to be higher than a much larger group, whose risks are more spread out and actuarial experience more reliable (NCSL, 2018). The ACA put together small groups in the SHOP (or a comparable state exchange) as if they were one greater risk pool, granting groups that buy health insurance through SHOP or state exchanges the same kind of group bargaining and purchasing power that larger businesses enjoyed exclusively, pre-ACA (NCSL, 2018).

However, our survey data indicates that since the ACA became effective, the SHOP exchange has been the least used by small groups with less than 50 FTEs. Only about 13 percent of them purchased from this exchange (KFF/HRET, 2015; KFF/HRET, 2016). Although rates were group-discounted, almost two-thirds of surveyed small business owners found the SHOP exchange plans either expensive or restrictive (e.g., at least 70 percent of small firm workers had to enroll in any of them). Two-thirds eventually found better plan rates elsewhere, by direct purchase, or through their brokers and agents (KFF/HRET, 2015; KFF/HRET, 2016). With risk

pools declining in size and mix in the next four years, the federally facilitated SHOP exchange, covering 33 states, closed by the end of 2017. Small firms were then directed to insurance carriers or to the remaining SHOP-registered agents and brokers (Jost, 2017). In most states, small firms also communicate and signal through these representations to pool together (i.e., through purchasing pools or CO-OPs) and buy SHOP plans and AHPs.

Direct purchases from insurance carriers entail transaction costs on the part of the employer. Health insurance markets in the U.S. are “characterized by imperfect information, complex products, and substantial search frictions” (Karaca-Mandic *et al.*, 2018). Particularly time-consuming and administratively demanding are search and information costs (e.g., comprehending and comparison-shopping for plans), as well as bargaining and decision costs, including matching plan offerings with employee demographics and preferences and making plan-related choices (Morrissey, 2020). As many as 87 percent of our surveyed small firms thus rely on commission-based brokers and agents to purchase, upgrade, or replace their health plans (Karaca-Mandic *et al.*, 2018).

Small group propensities

Firm size based on FTE count is a strong predictor of small group healthcare coverage ($r = 0.84$), as depicted in Table 1. The larger the workforce, the higher is the probability that a small business (<50 FTEs) will offer at least employee-only healthcare coverage. This positive correlation can be gleaned in both pre-ACA and ACA years. The average (x) percentage of sponsoring small firms over a five-year period (2015-2019) ranged from approximately 45 percent (i.e., with 3-9 employees) to 63 percent (10-24 employees) to 78 percent (25-49 employees). A similar finding pertains to small firms that are subject to the employer mandate (50-100 workers) but nonetheless benefit from certain statutory incentives (e.g., SHOP plans). Around 92 percent of businesses with at least 50 workers sponsored health insurance. Small businesses that do not offer it tend to be much smaller in terms of workforce size (<8 FTEs), rely more heavily on a part-time workforce, and/or tend to employ lower-income workers.

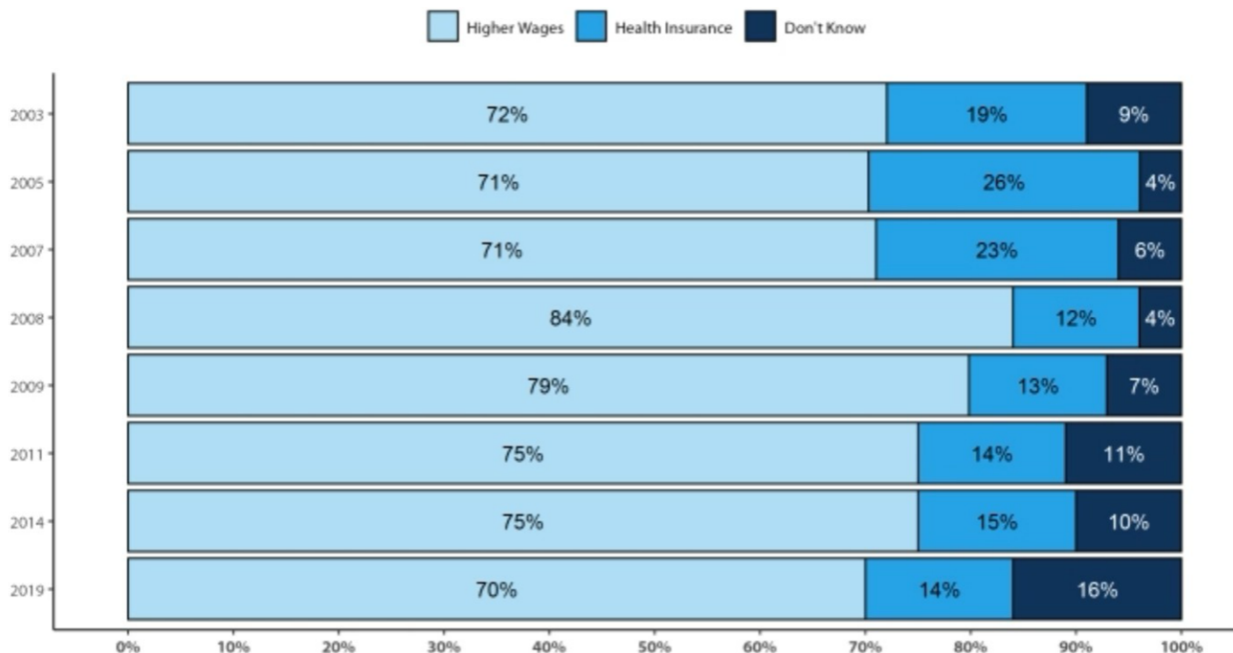
Because almost 9 out of 10 small business owners use brokers and agents to choose their employer-sponsored health plan design and benefits (KFF/HRET, 2015-2019), one could fairly assume the competitiveness of premiums and employee cost-sharing in small group plans (eHealth, 2018). On the other hand, from an organizational behavior perspective, the employer incentive differential stemming from premium affordability to, and desired benefits package of, the owner is suggested by the KFF-HRET data on the compensating wage differential in Figure 1. This refers to the minimum additional amount of income that workers will have to be offered in order to motivate them to accept or perform their jobs either in the absence, termination, or reduction of insurance coverage (Rosen, 1986; Kaufman, 2005). Asked in the KFF/HRET surveys whether they believe their insured employees would rather receive an additional \$2 per hour wage increase instead of healthcare coverage, as many as three-fourths of small business owners consistently responded in the affirmative over the years. Yet, these same owners are also very unlikely to opt for a wage differential which could be more expensive, as salary increases and inflation costs over time compound on the \$2 additional wage.

Table 1. Employer-sponsored health insurance by small firm size (average % of sponsoring firms)

Firm size (# employees)	2000 (x = 68)	2002 (x = 65)	2004 (x = 62)	2006 (x = 60)	2008 (x = 62)	2010 (x = 68)	2015 (x = 56)	2016 (x = 55)	2017 (x = 53)	2018 (x = 56)	2019 (x = 56)
<10	57	58	52	49	50	59*	47	46	40	47	47
10-24	80	70*	74	73	78	76	63	61	66	64	63
25-49	91	87	87	87	90*	92	82	80	78	71*	77
50-199	97	95	92	92	94	95	92	91	92	91	93

*Estimate is statistically different from the immediately preceding year, whether or not the preceding year is indicated in Table 1 ($p < .05$). Because KFF/HRET surveys collect information from a large sample of all U.S. firms, even seemingly large differences among them may not necessarily be statistically significant from year to year.

Figure 1. Small firm assessment of compensating wage differential*



*No statistically significant differences in percentage distributions ($p < .05$) were observed from year to year in the period covered.

Source: KFF (2019). *Employer health benefits: 2019 annual survey*. San Francisco, CA: KFF. Wage differential question asked of sampled small businesses/owners was whether they believe their employees would rather receive an additional \$2.00 per hour in salaries/wages in lieu of employer-sponsored health insurance. This question was asked in the KFF/HRET surveys only in years indicated in Figure 1.

Take-up rates among small and large firms are contrasted in Table 2. Take-up refers to the proportion of eligible employees who actually enroll in employer-sponsored health insurance. This excludes workers who opted out of obtaining coverage from their employer because federally subsidized insurance from the individual market and (broader) access to Medicaid under the ACA offered better alternatives to small group insurance.

Even among small firms, the five-year take-up rate is high enough, averaging over 75 percent. It is only slightly, albeit statistically, different ($r = 0.73$; $p = 0.44$) from the average take-up rate in large firms ($x = 78.2$ percent) during the same period. On the one hand, high take-up rates among small business workers anchors the propensity of small business owners to offer health insurance in place of a compensating wage differential. On the other hand, they appear to validate the value-added contribution of healthcare coverage to human resource goals, including employee attraction and recruitment, retention, and productivity (Chamberlain and Tian, 2016; AHIP, 2018).

Table 2. Eligibility and take-up rates (%) in small group health plans

Plan eligibility and take-up	2015		2016		2017		2018		2019	
	Small	Large	Small	Large	Small	Large	Small	Large	Small	Large
Employees eligible for employer health insurance	83*	79*	83	80	83*	80*	83*	79*	84*	80*
Eligible employees who participate in employer health insurance	76*	79*	77*	79*	76*	78*	74*	77*	73*	78*

*Eligibility or take-up rate is statistically different between small and large groups in the given year ($p < .05$).

Plan design and attributes

Which plan design/type and benefits small groups choose to offer and how much the employer and insured worker (enrollee) contribute typically result from back-and-forth communication with and negotiation through their brokers and agents (KFF/HRET, 2015-2019). Rarely are worker preferences surveyed actually by employers. But plan information is communicated to workers during recruitment, the open-enrollment period, and for qualifying life events.

Certain plan characteristics and attributes statistically differ between small and large businesses, offering decision-making insights into how small businesses finance and retain healthcare coverage, despite their substantial and rising costs.

Approximately 8 in 10 small groups in Table 3 offer only one type of plan, often a PPO

which enrolled around 46 percent of all small group workers between 2015 and 2019. In contrast, statistically higher numbers of large firms, averaging 75 percent from 2015 to 2019, offered a PPO. In a PPO, enrollees may seek care outside the provider network, but generally pay higher premiums and cost-sharing for network providers (e.g., copays, deductibles, co-insurance, maximum out-of-pocket cost). Large firms are equally likely to offer one (45 percent) or two (42 percent) plan types; they are more than four times likelier to offer three or more plan types than small firms.

Table 3. Small group healthcare coverage by plan design and attributes (%)

Plan structure	2015		2016		2017		2018		2019	
	Small	Large	Small	Large	Small	Large	Small	Large	Small	Large
<i>Plan types (#) offered:</i>										
—One	84*	52*	84*	47*	83*	45*	81*	42*	76*	39*
—Two	12*	39*	14*	37*	15*	42*	16*	45*	20*	47*
—Three or more	3*	9*	2*	16*	2*	13*	3*	13*	4*	14*
<i>Type of plan:</i>										
—Conventional indemnity	1	2	2	2	2	1	<1*	1*	3	1
—HMO	17	23	22	27	17*	31*	32	25	26	26
—PPO	49*	77*	34*	73*	49*	73*	49*	75*	48*	78*
—POS	27*	14*	33*	17*	28*	10*	14	11	24*	14*
—HDHP (may have HSA, HRA, other SOs)	25*	41*	27*	51*	23*	53*	27*	58*	27*	57*
<i>Plan coverage:</i>										
—Single coverage only	2	0	11	0	5	0	5	0	6	0
—Spouse	98	100	89	99	94	100	97	99	94	99
—Children and other dependents	96	100	88	100	92	100	95	100	94	100

*Percentage distribution is statistically different between small and large groups for the given year ($p < .05$).

HDHPs (high deductible health plans) ranked second among the plan offerings of a large firm (52 percent). A plan is considered an HDHP if the single (individual) and family deductibles met or exceeded the IRS-prescribed annual deductible (\$1,650 for single coverage and \$3,300 for family coverage in 2025). Traditionally paired with health savings accounts (HSAs), HDHPs have much lower premiums than a fully-insured plan (e.g., PPO, HMO, POS, EPO). The major disadvantage of an HDHP is the generally staggering out-of-pocket costs to

the enrollee and their family because the annual deductible is also way higher (Morrisey, 2020). This financial risk to the insured, along with lower average wages in small firms, suggest why only one in four small firms offered an HDHP in Table 3. Cost-sharing in small businesses, particularly for in-network providers and services, tends to be lower than in large group HDHPs. The same strategic risk-aversion is observed among small firms that adopted a self-insured/funded plan which transfers risk exposure directly to the employer (Park, 2000; Morrisey, 2020). An average of only 15 percent of small groups in Table 3 chose to self-insure compared to over 62 percent of large groups where it is the predominant plan design (KFF/HRET, 2015-2019).

In terms of actuarial value, nearly or slightly over one-half of small firms tend to offer a silver-tier plan or its equivalent (KFF/HRET, 2015-2019). This carries a 70 percent actuarial value, which is the portion of medical expenses that a health plan covers on average compared to the enrollee's cost-sharing responsibility (in this case, 30 percent). A silver plan has lower monthly premiums, but enrollee cost-sharing is higher than the gold and platinum plans which respectively carry 80 and 90 percent actuarial value. Most large firms prefer to offer a gold-tier plan or its equivalent.

The number of covered lives may provide yet another indication of cost-containment among small firms. Despite significant variance from year to year in Table 3 (minimum = 2 percent; maximum = 11 percent), some small firms offer employee coverage only, while none of the large firms do so. Spouses, children, and other dependents (e.g., same-sex partner) can be extended coverage in nine out of 10 small group health plans, and practically 100 percent of large group insurance, although differences in some years were statistically insignificant.

Plan financing

How much and in what ways will a small business finance employee healthcare coverage are usually owner and/or management decisions which underpin plan design. Key considerations include the *pure premium* and the *loading fee* or expense ratio covering insurer profit/mark-up for objective risk, marketing costs, benefits coordination, claims processing and adjudication, and other administrative overheads (Green, 2021). The reviewed literature is helpful in pointing out that such decisions are not exclusively budget-driven but linked to employee recruitment, retention, and productivity, especially considering that employee turnover in small groups is generally higher (Mendoza, 2024).

Table 4 shows that annual premium rates for employee-only coverage in small groups averaged \$6,622 for five years, which are not statistically much different from large groups ($x = \$6,722$). Average premium growth rate is around 4 percent annually for small and large groups alike. But the financial burden on small firms is much heavier, with single premium rates constantly rising (or skyrocketing). An eHealth survey, for example, reported that 78 percent of small business owners were either somewhat or very concerned about being able to continue paying for employee healthcare due to increasing rates. About 62 percent indicated that a 15 percent premium increase would be unaffordable, forcing them to drop or reduce coverage or shop for a new plan (eHealth, 2018). These lend empirical support to the employer's cost-calculus between insurance affordability and strategic human resource objectives.

Table 4. Average annual premiums, small group health insurance, all plan types

Premium attribute	2015		2016		2017		2018		2019	
	Small	Large	Small	Large	Small	Large	Small	Large	Small	Large
<i>Average total premium (\$)</i>										
—Single coverage	6,163	6,289	6,429	6,438	6,486	6,776	6,814	6,930	7,218	7,175
—Family coverage	16,625*	17,938*	17,546*	18,395*	17,615*	19,235*	18,739*	19,972*	20,236	20,717
<i>Average total premium by firm ownership (\$)</i>										
—Private for-profit	5,841**	5,934**	6,164**	6,021**	6,204**	6,339**	6,659*	6,665*	7,010*	6,714*
—Private not-for-profit	7,039**	6,829**	7,124**	6,863**	7,194**	7,061**	7,213*	7,281*	7,406*	7,669*
—Public	6,962**	6,860**	7,543**	7,060**	7,350**	7,549**	6,964*	7,387*	8,587*	7,923*
<i>Average employee premium (\$ contribution (% total premium))</i>										
—Single coverage	899(15)	1,146(18)	1,021(16)	1,177(18)	1,030(16)	1,290(19)	1,133(17)	1,207(17)	1,035(14)	1,330(19)
—Family coverage	5,904(36)*	4,549(25)*	6,597(38)*	4,719(26)*	6,814(39)*	5,264(27)*	6,782(36)*	5,046(25)*	7,805(39)	5,271(25)
<i>Average employer premium (\$ contribution (% total premium))</i>										
—Single coverage	5,264(85)	5,142(82)	5,408(84)	5,261(82)	5,456(84)	5,486(81)	5,681(83)	5,723(83)	6,183(86)	5,845(81)
—Family coverage	10,720(64)*	13,390(75)*	10,949(62)*	13,676(74)*	10,801(61)*	13,971(73)*	11,957(64)*	14,926(75)*	12,431(61)	15,446(75)

*Percentage distributions are statistically different (horizontally) between small and large groups in terms of corresponding premium-related attribute ($p < .05$).

**Percentage distributions are statistically different (vertically) only by firm ownership category within small or large groups, but not (horizontally) between any other premium-related attributes ($p < .05$).

It is in the family coverage premiums where small and large businesses statistically differ in Table 4. The five-year average is lower among small firms (\$18,152) in contrast to large ones (\$19,251), although both increased yearly at an average of 5 percent. Since small group health insurance tend to have higher loading fees, particularly administrative and marketing costs (Long, Rae & Claxton, 2016), lower family premiums in small firms implies that healthcare benefits are also less generous on average. Offering affordable family coverage, even for less generous benefits, could nonetheless help attract and retain needed talent in small businesses.

Table 4 also shows that small and large employer contributions to single plan premiums do not statistically vary, both averaging slightly over 80 percent from 2015 to 2019. Most enrollees in small firms contribute approximately 16 percent of total premium cost. However, small and large firms vary with respect to family plan contributions. Small employer contributions were consistently around 60 percent of premium cost ($x = 62$ percent). Large

employer contributions were consistently at a much higher rate, in the mid-70s ($x = 74$ percent) during those five years. Thus, small group workers pay 12 percent more in family premiums ($x = 38$ percent) than their counterparts in large groups. The difference likely accounts for greater budgetary and other financial constraints faced by small businesses in covering employee spouses and dependents.

Employee-paid premiums statistically vary once stratified by group ownership in Table 4. Higher premiums are consistently charged by public ($x = \$7,481$) and private not-for-profit ($\$7,195$) firms compared to private for-profit firms ($\$6,376$) in either small or large group markets. Pricing differences suggest more generous benefits offered by public and private not-for-profit health plans.

Risk management

Health insurers resort to favorable selection, HMO-style disincentives, utilization review, and other enrollee-focused approaches to reduce risks of adverse selection and healthcare over-utilization or moral hazard by the enrollee (Morrisey, Kilgore *et al.*, 2013; Morrisey, 2020). Small firms have less options available in containing the financial burden or implications of enrollee behaviors once these small firms elect to offer coverage (Morrisey, 2020), but our findings suggest some approaches might be strategically advantageous.

For one, Table 5 indicates that small groups are more predisposed to offer healthcare coverage with higher deductibles than large groups. This means enrollees pay 100 percent of covered treatments and services until a stipulated dollar amount is reached at which time the group health plan starts to pay while enrollee is charged coinsurance. That dollar amount is generally higher in small groups. Even if there are fewer small group HDHPs, the number of their workers in single coverage HDHPs with deductibles of \$2,000 or more is typically higher than large groups. Average single coverage enrollment in these plans for small and large groups, respectively, are also statistically significant from 2015-2019:

\$1,000 deductible: 64 percent; 46 percent
 \$2,000 deductible: 42 percent; 17 percent
 \$3,000 deductible: 23 percent; 7 percent

Generally, the higher a plan's deductible, the lower the premium cost, as more risk is ostensibly shifted to the enrollee who exposes more “skin in the game” (LaMontagne, 2014; (Agarwal et al., 2017). In this sense, offering health insurance becomes an employer or business owner balancing act between deductibles and premiums. Because small groups pay the majority of premium costs in both single and family plans (see Table 4), the employer (and employee) generates premium-savings from higher deductibles, which in Table 5 are sometimes at levels usually unseen in large group plans for the years covered in this study (e.g., \$3,000). Combined with the small group’s propensity to offer only one plan or one plan design/type (Table 3), health plans with higher deductibles also lead to administrative and other transaction cost-savings to small groups. The cost of claims processing and adjudication alone decline with reduced healthcare access (Green, 2021). Yet, there is no gainsaying that the effectiveness of higher

deductibles can hinge on whether workers in a small group are predominantly low-income, younger or older, healthy or living with a chronic condition, etc. (Agarwal et al., 2017), considering that the primary goals of small group insurance relate to job satisfaction, especially after it has produced an incentive differential to the owner or employer.

Table 5. Cost/risk management approaches among small firms offering health insurance (%)

Cost/risk reduction	2015		2016		2017		2018		2019	
	Small	Large	Small	Large	Small	Large	Small	Large	Small	Large
Insured employees in (single) plans w/ high annual deductible:										
≥\$1,000	63*	36*	65*	41*	58*	48*	68*	54*	68*	50*
≥\$2,000	39*	12*	45*	16*	37*	15*	42*	20*	45*	22*
≥\$3,000	—	—	—	—	22*	6*	22*	7*	24*	9*
Firms adopting narrow(er) network plan/s**										
	8	9	7	6	8*	11*	6*	10*	5*	11*
Firms eliminating hospitals/healthcare systems from their networks in past year**										
	10	6	6	5	6	4	2*	5*	1	3
Firms that searched for new plan/carrier in past year:										
— Shopped for new plan/carrier	47*	38*	52*	35*	59*	37*	62*	50*	53	49
— Shopped + changed plan/carrier	24	24	21	27	29	26	26*	20*	18	22
Firms offering incentives to employees enrolling in spousal plan										
	9	11	10	12	15	13	13	12	—	—
Firms w/ incentives to employees for not enrolling in employer plan (if w/ out spouse/spousal plan enrollment)										
	7	5	9*	15*	18	17	16	15	—	—
Non-offering firms providing funds/reimbursements for employees to buy insurance on their own***										
	17	—	11	—	16	—	9	—	11	—
Non-offering firms that terminated health insurance***										
	25	—	19	—	13	—	20	—	13	—

*Percentage distribution is statistically different between small and large groups ($p < .05$).

**Around 5% of small groups on average considered either a narrow(er) plan network (< providers) or eliminating hospitals and healthcare systems from their networks in the past year to reduce cost.

***Refers to small groups that either do not offer group health insurance, or have terminated it.

—Not surveyed/reported by KFF/HRET during the corresponding year.

Table 5 suggests that small firms tend to communicate more frequently with brokers and agents during the year to *shop* for new health insurance and/or carrier offering lower premiums and deductibles and/or broader provider networks. Statistically significant for all five years, except 2019, 50 percent or more of small firms shopped for replacement plans and/or carriers to contain cost, as opposed to less than 40 percent of large groups. Other risk mitigation strategies

sometimes used by small firms, but generally *not* any statistically different from large firms, include offering narrow(er) provider networks, eliminating certain hospitals and healthcare systems from their current network, and terminating their current health plan and/or carrier. In this sense, brokers and agents help offer lower premiums, especially in more competitive broker/agent markets. Premiums are also less dispersed in these markets (Karaca-Mandic *et al.*, 2018).

Table 6 indicates that many small and large firms alike choose to set up risk management programs or options within their health plans to help enrollees identify health issues and manage chronic conditions. Plan offerings include health risk assessments, biometric screenings, and health promotion programs (Mendoza, 2025). Like large firms, too, small groups do not typically offer employee incentives either to enroll in spousal plans or non-employer plans in lieu of participating in their employer-sponsored health insurance.

Table 6. Risk management programs/options offered in small and large group health plans (% all plan sponsors)

Year	Programs to help stop smoking	Programs to help lose weight	Other lifestyle/behavioral coaching	Disease mgmt. programs	At least 1 other wellness program*	Penalty for tobacco use or vaping
2015	42	40	40	33	50	-
2016	38	35	38	-	47	-
2017	41	35	47	41	59	16
2018	45	43	40	-	54	17**
2019	37	32	40	29	51	-

*Most common examples include: classes in nutrition and healthy living, gym membership discounts/credits, on-site exercise facilities, and wellness benefits offered to spouses/dependents.

**Only large firms of ≥ 200 workers were surveyed.

– Not surveyed/reported by KFF/HRET during the corresponding year.

The lure of group association

When healthcare coverage costs increase and become challenging to small businesses or comparison-shopping yields less costly options, Table 5 suggests that small businesses rarely consider, or at least implement, alternatives to fully-funded plans (13 percent), including self-funding, unlike their large business counterparts. Brokers and agents are also unlikely to suggest self- or partial-funding, given their fee-based incentives. Group insurance still appears to be

more cost-efficient and attractive even to small business employees (Freedman, 2020). Instead, association health plans (AHPs) might offer a more viable alternative (Mendoza, 2024). For that reason, we explore in this section the strategic advantages and limitations of forming AHPs.

The ACA and other federal regulations allow small groups to band together and form an “association” to purchase fully-funded health insurance through an insurer-issued AHP, unless they opt to associate to self-fund their AHP. Primarily intended to encourage small groups to sponsor healthcare coverage that might otherwise be prohibitively expensive if separately purchased from the small group market, AHPs can better pool similar risk groups (e.g., various small realty firms or bakeshops, rather than one single realty firm or bakeshop). AHPs thus gain better leverage and can take advantage of greater cost-savings typically available only to large group plans. AHPs also tend to “have lower administrative costs, perhaps because of [their] ongoing relationship and communication with members” (Morrisey, 2020: 366).

Certain ACA healthcare reforms apply only to small group (and non-group) insurance plans to protect them from their inherent disadvantages in pricing and coverage. After all, insurers have little incentive to offer — and small firms and individual purchasers have little leverage to negotiate — comprehensive benefits, in light of the higher actuarial risks of insuring a smaller pool of workers (Chaikind *et al.*, 2010). As noted earlier, these mandatory protections and standards include the 10 essential health benefits that plans must offer; tiered actuarial values that specify policy-holder and insurer cost-sharing for covered benefits; modified community rating method that allows premiums to vary only by age, tobacco use, geographic location, and family size; single risk-pooling so that underwriting is based solely on the pooled health risks of policy-holders; and “risk adjustment” by which health plans that disproportionately enroll higher-risk people receive payments from plans that enroll younger and healthier individuals. These reforms are costly to many small group and non-group insurers (Mendoza, 2017). The ACA drafters did not impose them on fully-insured and self-insured large group plans because “they felt that these plans’ covered benefits... were as good if not better” than what the ACA prescribed for the small group and non-group health plans (Condeluci, 2018: 6). There was likewise the presumption that large groups are in a better position to negotiate and bargain with insurers, and therefore, do not need as many protections (Condeluci, 2018).

Interviews with AHP Inc., an interest group representing various AHPs, underscore two key aspects of group associations offering AHPs in lieu of traditional small group plans. The first is legitimacy of organizational scope. Commonality of interest (small group members belong to the same trade, industry, line of business, or profession) is required by U.S. Department of Health and Human Services (HHS) implementing regulations. Another issue that AHPs, especially newly set up ones, often need to contend with at the state level is a “seasoning requirement.” The association (not the small group) sponsoring the AHP must be of minimum age (often three years old) to discourage “fly-by-night” associations and fraudulent carriers, brokers, and agents (Coleman, 2024).

Efficient communication, especially in terms of signaling, is essential to small group association. These groups do not need to know one another prior to founding an AHP. Brokers and agents with AHP experience are hired do the needed groundwork and communications. Some small groups, usually with existing multiple business relationships, prefer to do it on their own. In this *de novo* scenario, those sharing an acceptable commonality of interest communicate

and meet up to perform several tasks: 1) Establish a formal sponsoring association with a governing body (e.g. directors or trustees) and by-laws; 2) Provide a defined structure for leadership roles (e.g. president, treasurer, etc.); and 3) conduct and report on regular meetings. Whichever route the association takes, there must be compliance with federal and state regulations, for which they may approach outside counsel with AHP expertise (Coleman, 2024).

However, AHPs are not without controversy. A 2011 HHS regulation practically disallowed small groups (with an aggregate total of over 49 or over 100 FTEs, depending on the state) from sponsoring a fully insured large group AHP after informing their insurers of their workforce size (HHS, 2011). This came amid fears that many small groups were fleeing the small group market for the AHP market, and were seeking equal treatment as a large group to contain skyrocketing premiums and operational costs without having to offer “just as good, if not better” standards and protections mandated by the ACA. AHPs that are treated as a large group plan tend to generate cost-savings to employer and employee alike better than a small group AHP by “negotiating [more] leverage based on economies of scale” and offering an even “bigger ‘risk pool’” (Condeluci, 2018, p. 9). The 2011 HHS regulation led to a drastic decline in fully-insured AHPs, after many of small firms chose to discontinue their AHPs amid the risk of non-compliance (Condeluci, 2018).

The first Trump administration sought to reverse the decline and directed the U.S. Department of Labor (DOL) — tasked with administering employee benefits law — to expand AHP availability (Executive Order 13813, 2017). The DOL did so by considerably loosening the definition (or interpretation) of a qualified “*bona fide group or association of employers*” under the 2011 HHS regulation. By 2018, the DOL had released a highly contentious “AHP rule” (known as Pathway 2), allowing small group associations to pass the commonality of interest test if: 1) member-groups are in the same trade, industry, line of business, or profession, regardless of geographic location; *or* 2) member-groups have their principal place of business in one state or metropolitan area (spanning a tri-state area), whether or not it spans more than one state (DOL, 2018). Removing the strict geographic limitation through the first criterion was “critical for national trade associations, franchisees, and companies with “cooperative” members... [and which have for decades] wanted to offer some type of health coverage to their [small] group members through a fully-insured or self-insured AHP *on a nationwide, or a regional basis*” (Condeluci, 2018: 9). The second criterion allows “unrelated” employers from different trades, industries, businesses, and professions (e.g., chambers of commerce), meeting the needed FTE aggregate counts, to sponsor an AHP, but only if they are principally located in the same state or metropolitan area. A major objection to Pathway 2 stemmed from the fear that associations based on different geographies and/or unrelated small businesses would expand AHP market share and draw people out of ACA risk pools, costing small group insurers more and raising small group premiums even more.

The DOL’s Pathway 2 went one step further by permitting self-employed individuals without employees (referred to as “working owners”) to act both as an employer sponsoring a group health plan and an employee participating in an AHP (DOL, 2018). These working owners could then participate in an AHP established by other “related” small groups (e.g., solo realtors who belong to a realty association) in the same industry/profession, even if they are located in different geographic areas. This provoked further opposition.

Several consumer advocacy groups, insurers, and state governments challenged small group formation of large group AHPs through federal litigation because of their exemption from ACA mandatory standards and protections for small group health plans. Consider the opposition of the American Medical Association and seven other physician organizations. They asserted that Pathway 2 serves to lure small groups into AHPs until a medical emergency strikes. It is then that they discover essential care is not covered, which is particularly devastating to economically vulnerable and unhealthy populations and their dependents (Golder, 2019). Others claim that Pathway 2 insurers can easily game the system through higher cost-sharing, narrower provider networks, and reduced coverages for large group AHP enrollees (Condeluci, 2018), thereby turning “associations as a vehicle for bootstrapping them into the large-group market” (Golder, 2019: 25).

AHP demographics

Pathway 2 led to the rapid growth of AHPs starting in 2018. It allowed 3.2 million small groups to set up the equivalent of large group AHPs for their almost 50 million employees, and 27 million independent contractors to join an AHP for the first time (Cassidy, 2024). The 2019 AHP demographic survey provide context to the impugned DOL rule.

As much as 71 percent of 28 new AHPs since 2018 were set up by regional associations of small groups. Another six AHPs were in the process of forming by 2019. The vast majority of these 34 associations shown in Table 7 were chambers of commerce, some of which were multiple chambers sponsoring a single health plan (e.g., all 29 chambers in Vermont). Most benefitted from Pathway 2’s second criterion that allows small groups in different industries and professions (i.e., “unrelated” employers) but located in the same geography to form AHPs. Non-chamber regional associations included business groups (e.g., Small Business Association of Michigan), regional trade groups (Wisconsin Manufactures & Commerce), and philanthropic organizations (Business Fund for Texas Children Battling Cancer). On the other hand, multi-state professional AHPs, which limit membership to small groups within the same line of work/trade, took longer to reach the association market because they required more preparation and state filings. AHPs formed under the 2011 HHS regulation (i.e., prior to Pathway 2) are comparably much fewer in number (AHP.com & Coleman, 2023a).

Over 86 percent of the new AHPs in Table 7 were fully insured, rather than self-funded, as a large group plan. Preference for full insurance owed largely to AHP reliance on an insurer with greater capital resources to withstand catastrophic medical expenses (Morrisey, 2020). UnitedHealthcare and Blue Cross/Blue Shield were the most frequently chosen carriers. Fully-insured AHPs typically adopted strategies to mitigate adverse selection and promote financial stability, including a set minimum number of covered FTEs within the range allowed by the state in which the AHP resides, minimum period for association membership of a small group before accessing an AHP, and a wait time prior to small group withdrawal from the association (AHP.com & Coleman, 2023b).

Table 7. Demographics of small firm associations/AHPs formed under Pathway 2 (n = 34)

Association demographics	<i>f</i> (2018-19)
<i>Associations formed</i>	
–Active	28
–Inactive (i.e., under formation)	6
<i>Commonality of interest</i>	
–Regional* (same geography*; may be different lines of work/trade)	20
–Professional (only same line of work/trade; may be different geography)	8
<i>Minimum group size for association membership</i>	
–One (“working owner”)	12
–Two to 50 FTEs	15
–51 to 100 FTEs	1
<i>AHP funding</i>	
–Fully funded/insured	24
–Self-funded/insured	4
–Partially-funded/insured	0
–Inactive but leaning toward fully-funded/insured	6
<i>Plan types/options available (average #)</i>	
–Fully-funded/insured	11
–Self-funded/insured	11
–Partially funded/insured	0
<i>Benefits covered (essential care)</i>	
–Comprehensive**	28
–Limited/narrow	0
<i>Medical savings account options (e.g., HSA, HRA)</i>	
–Offered (at least one)	14
–None offered	14

*Approximately 4 of 5 regional associations formed under Pathway 2 were sponsored by (single or multi-) chambers of commerce. Geographic location refers to municipality, city, county, regional, or state..

**Minimum coverage includes doctor and specialist visits, hospitalization, emergency care, prescription drugs, maternity care, preventive care, and mental health care. Lack of coverage generally observed for pediatric dental care and pediatric vision care, but these benefits are not mandated by law in large group plans and may be offered through stand-alone dental and vision plans.

Although plan design and benefits were fairly similar to a traditional small group plan, cost-savings to large group-styled AHPs came primarily from premium rates that were at least 30 percent lower than those for ACA-regulated non-group and small group plans (CBO, 2019). In terms of scope of benefits, most of the new AHPs offered comprehensive coverage that resembled mandatory essential health benefits for non-group and small group plans. These included primary care physician and specialist visits, hospitalization, emergency care, prescription drug coverage, maternity coverage, preventive care, and mental health care. Where large group AHPs largely differed from ACA standards appears to be in (their lack of) pediatric dental care and pediatric vision care coverage, which might be offered through stand-alone dental and vision plans.

Half of the new AHPs additionally provided pretax savings account options, especially a health savings account, which the enrollee could use for out-of-pocket costs, including copays, deductibles, and coinsurance (AHP.com & Coleman, 2023a).

Half of these AHPs were also restricted in membership to small groups with two to 50 FTEs. Less than half (43 percent) were open to the self-employed (AHP.com & Coleman, 2023a). A total of four million workers were estimated to have enrolled in AHPs as of 2018 as a result of Pathway 2 (DOL, 2018).

However, the Biden presidential administration rescinded Pathway 2 in 2024 after federal litigation ensued, effectively returning AHPs to pre-2018 (HHS) regulatory guidance. There is talk of reviving it in some shape or form under the second Trump administration (2025-present). Nonetheless, several large group AHPs have either adopted small group standards and protections under the ACA or have individually insured their member-businesses as a small group while gaining for them AHP-like lower premiums and other cost-savings.

4. DISCUSSION AND CONCLUSION

Small business research into the health insurance market is academically compelling in light of the dearth of literature which essentially focuses on incentives and challenges to employer sponsorship of employee healthcare coverage. It is also insightful from a theoretical and practical business standpoint. We sought in this study to discover why (and how) small firms or groups as well as associations of these firms or groups manage to offer healthcare coverage in pursuit of human resource objectives and despite statutory exemptions and disincentives. We did so by inquiring into how small groups might overcome the substantial cost burden of plan sponsorship.

Our cross-sectional analysis of the EHB-AS dataset reveals that small business decision-making is more nuanced than the primordial objective of employee job satisfaction relative to recruitment, retention, and development of needed skill and talent. Doubtless, statutory incentives to small group healthcare coverage are in place. But they can be offset by restrictive or burdensome qualification requirements, leading brokers and agents to assume an important role in small group plan choices and alternatives. Because almost nine out of every 10 small firms employ them, premium rates can be quite competitive. Yet, no matter how competitive these rates may be, they also have to be affordable relative to firm size and income and employee demographics. After all, the majority of premium costs for both single and family plans are

employer-paid, besides the added operational costs that these small employers have to bear.

In this regard, we posit the view that an incentive differential equally plays an important role in steering small firms toward human resource objectives. To the extent that premium rates are deemed affordable by a small business, and/or group coverage makes the business owner better off than getting it elsewhere or remaining uninsured, it is likely that the small firm will sponsor health insurance. That is to say that cost-efficiency to a small firm and its owner (the employer differential) go hand in hand with the expected gains of employee job satisfaction in motivating small firm sponsorship notwithstanding the limited advantages of statutory incentives and the substantial costs of premiums and plan administration. Otherwise, the owner's cost-calculus might favor a compensating wage differential. Three-fourths of small business owners, in fact, believe that their workers would rather receive a higher wage than healthcare coverage.

With our findings, we distinguished organizational problem-solving and decision-making toward cost-efficiency within and between small firms. Doubtless, employees who enroll within their organization for small group health insurance have considerably more consumer protections, and their insurers have additional standards to meet, under the ACA. But there are inherent exposure risks to small group plan sponsors. We discovered strategic differences between plans offered by small and large firms through which small firms seek to mitigate risks (and therefore costs) based on employer propensities, plan design and benefits, financing, and risk management. About eight in 10 small firms choose to offer only one type of plan, usually a PPO. In contrast, large firms are likely to offer two or more plans and plan types. Smaller firms are more likely to offer a silver-tier plan which charges lower premiums but more in terms of employee cost-sharing. Some small firms offer only single employee-coverage plans, while none of the surveyed large firms do. Small firms also tend to subsidize family plans (if offered) at a lower rate, while their lower family premiums imply less generous health benefits on average. Small firms are much more likely to use fully-funded plans with higher deductibles to contain administrative overhead and transaction costs. And through brokers and agents, small firms tend to shop for new or better health insurance and/or carriers more often due to skyrocketing premiums. Premises considered, we generally find small firms to be more proactive in utilizing certain cost or risk mitigation approaches than their large counterparts.

Besides these approaches, we find communication through signaling an important aspect of small group provision. Within a small firm exists various levels, channels, and forms of communication: between owners and management in selecting healthcare coverage and making funding decisions; between plan sponsors and their brokers and agents in plan selection and price negotiation; between employers and employees, particularly during recruitment, open-enrollment, and qualifying life changes; and between plan sponsor, insurer, and government in terms of regulatory compliance. Efficient communication based on market signals provides information about options and alternatives to employer and employee alike relative to premium pricing and cost-sharing, design and scope of benefits, quality and levels of coverage, plan administration, claims adjudication, and legal and regulatory compliance, among others.

Communication between small firms (and working owners) is also crucial in deciding whether it is more cost-efficient to band together and establish AHPs. The associational route has helped small firms contain their financial and operational costs as a result of better predictability

of medical claims due to wider actuarial experience, greater leverage from higher insurer revenue generated from associational membership, and lower administrative expenses relative to volume of association enrollees. Many small groups had taken advantage of rather conflicting federal regulations that allowed them to set up large group AHPs prior to the ensuing federal litigation in 2019. On the one hand, the DOL's Pathway 2 permitted even unrelated and geographically distant small groups to associate and form AHPs which critics contended undermined commonalities of interest and reduced ACA risk pools needed for better experience rating and reasonable premium setting. On the other hand, combining the number of FTEs in associated small groups to gain insurer treatment as a large group plan could cause AHPs to lose (or evade) the enrollee protections mandated by the ACA to safeguard the market vulnerabilities of small firms and their comparative disadvantages relative to large firms.

Within and between small businesses, we find transaction cost reduction essential in administering and retaining employee healthcare coverage. Transaction costs pertain to those incurred in "doing business in the market," including search and information, bargaining and decision-making, and monitoring and compliance, in contrast to the costs of production and delivery (Williamson, 1979). Search and information alone can be quite burdensome, even with the use of brokers, agents, third-party administrators, and legal counsel for plan selection, administration, and pricing. Small firms have more limited income and other resources compared to large firms and many funding decisions about healthcare coverage bear directly on organizational viability. Lacking in risk pool size, actuarial experience, and negotiating leverage, small firms generally face higher bargaining and decision costs, while insurers have every incentive to raise premiums and cost-sharing, limit ("narrow down") provider networks, offer minimum or reduced benefits, etc., to minimize risk exposure in small group plans. Legal and regulatory compliance and monitoring can be just as challenging. Complex or conflicting federal and state regulations apply and are often uncoordinated, as exemplified by the AHP route. While ACA enacted small group healthcare reforms, insurance is under state jurisdiction and certain federal regulations might be implemented differently. The legal controversies surrounding small group associations and increasingly attractive large group-styled AHPs is once again illustrative. Therein also lies the paradox: The more flexible and convenient it becomes for small groups to associate and set up large group AHPs, the greater is the likelihood of undermining small group protections and standards under the ACA. As the DOL conceded, increased small group leniency under Pathway 2 introduces increased opportunities for mismanagement or abuse, in turn increasing oversight demands on federal and state regulators (DOL, 2018).

We thus find actuarial (or financial) risk helpful but rather narrow as an analytical framework for small group problem-solving and decision-making in health insurance. Because transaction costs abound, and are more challenging to small group healthcare coverage, transaction cost analysis is well suited in explaining why cost-inefficiencies arise and which plan design, options, and transactional arrangements might better minimize them.

Finally, whether it makes sense to retain, enhance, or replace, if not terminate, employer-sponsored health insurance will also depend on facts and circumstances attendant in a small firm. Business risk will vary by small firm depending on employee demographics, including age and wage distribution, firm size and productivity, budgetary and related considerations in employer

plan sponsorship, and competitiveness of insurance options and pricing. There are also trade-offs to consider in plan choices, alternatives, and approaches. A plan sponsor, for instance, may succeed in reducing cost with higher deductibles or through self-funding, but in a way that discourages take-up and increases employee preference for a compensating wage differential which could prove disadvantageous to small business owners. Be that as it may, this study has drawn attention to certain behavioral propensities of and strategic approaches available to small firms, and their business implications, which may be useful and timely in mitigating costs, risks, and trade-offs in employee healthcare coverage.

CONFLICT OF INTEREST

The author has no relevant affiliations or financial involvement with any organization or entity with a financial interest in, or financial conflict with, the subject matter or materials discussed in the manuscript. This may include employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

ACKNOWLEDGMENT

The author acknowledges with thanks the helpful comments and suggestions of this journal's three anonymous reviewers and Kev Coleman of AHP, Inc., and the invaluable research and editorial assistance provided by Mary S. Danforth. As with any work of this nature, the usual caveat applies.

REFERENCES

- Agarwal R, Mazurenko O, Menachemi N (2017). High-Deductible Health Plans Reduce Health Care Cost And Utilization, Including Use Of Needed Preventive Services. *Health Affairs*, 36(10): 1762-1768.
- Agency for Healthcare Research and Quality (AHRQ) (2020). *Medical Expenditure Panel Survey Insurance Component 2019 Chartbook*. Rockville, MD: AHRQ.
- America's Health Insurance Plans (AHIP) (2018). *The Value of Employer-Provided Coverage*. Washington, DC: Luntz Global.
- America's Health Insurance Plans (AHIP) (2019). *Health Insurance: State-to-State 2019*. Washington, DC: AHIP Center for Policy and Research. Available at: https://www.ahip.org/wp-content/uploads/Nationwide_StateDataBook_2019.pdf.
- [AssociationHealthPlans.com](https://www.associationhealthplans.com) ([AHP.com](https://www.ahp.com)), Coleman K (2023a). First phase of new association health plans reveal promising trends. Available at: https://www.associationhealthplans.com/reports/new-ahp-study/#_edn6
- [AssociationHealthPlans.com](https://www.associationhealthplans.com) ([AHP.com](https://www.ahp.com)), Coleman K (2023b). Study: Chamber of commerce association health plans building political bridges while refuting opposition. Available at: <https://www.associationhealthplans.com/reports/chamber-of-commerce-ahp/>

- Chaikind H, Fernandez B, Peterson CL, Newsom M, Mulvey J (2010). Private health insurance provisions in Senate-Passed H.R. 3590, The Patient Protection and Affordable Care Act. Washington D.C.: Congressional Research Service.
- Cassidy B (2024). Letter to Julie Su, Acting Secretary, U.S. Department of Labor, February 29. Available at: <https://www.help.senate.gov/ranking/newsroom/press/ranking-member-cassidy-rebukes-proposed-dol-rule-restricting-health-coverage-for-millions-of-american-workers?fbclid=IwAR2OQeVMzUtl7XLUmXeJJkn3SzaXc9oTxf3c4qm2cksqoa4E5v5stbU7g>
- Chamberlain A, Tian G (2016). Which Benefits Drive Employee Satisfaction? *Glassdoor Economic Research*. Available from: <https://www.glassdoor.com/research/benefits-drive-employee-satisfaction/#>.
- Coleman K (2018). *Association health plans & the future of American health insurance*. San Francisco, CA: Publisher Services Publication.
- Coleman K (2024). Personal communication, July 26.
- Condeluci CE (2018). *Testimony before the Committee on Education and the Workforce, Subcommittee on Health, Employment, Labor, and Pensions*. Washington D.C.: U.S. House of Representatives. Available at: https://edworkforce.house.gov/uploadedfiles/testimony_condeluci_revised_3.20.18.pdf
- eHealth, Inc. (2018). *Small Business Health Insurance Costs, Trends and Insights 2017*. Santa Clara, CA: eHealth, Inc.
- Freedman M (2020). What Are the Benefits of a Section 125 Plan? *business.com*. Available from: <https://www.business.com/articles/benefits-of-section-125-plan/>.
- Golder C (2019). Brief of *amici curiae* American Medical Association, Medical Society of the State of New York, American Academy of Family Physicians, American Academy of Pediatrics, American College of Physicians, American College of Emergency Physicians, American College of Obstetricians & Gynecologists, and American Psychiatric Association in support of appellees. Available at: https://litigationtracker.law.georgetown.edu/wp-content/uploads/2023/04/NY_20190722_AMICUS-BRIEF-filed-by-American-Academy-of-Family-Physicians-et-al.pdf
- Green M (2021). *Understanding Health Insurance: A Guide to Billing and Reimbursement*. Boston, MA: Cengage Learning.
- Hall MA, McCue, MJ (2018). The Health of the Small-Group Insurance Market (The Commonwealth Fund). Available from: <https://doi.org/10.26099/t0k3-f906>.
- Jost T (2017). CMS Announces Plans To Effectively End The SHOP Exchange. *Health Affairs*. Available from: <https://www.healthaffairs.org/doi/10.1377/hblog20170515.060112/full/>.
- The Kaiser Family Foundation and Health Research & Educational Trust (KFF/HRET). *Employer Health Benefits, Annual Survey* (for the years 2015-2019). San Francisco, CA: Henry J. Kaiser Family Foundation.
- The Kaiser Family Foundation (KFF) (2019). Health Insurance Coverage of the Total

- Population. Available from: <https://www.kff.org/other/state-indicator/total-population>.
- Karaca-Mandic P, Feldman R, Graven P (2018). The Role of Agents and Brokers in the Market for Health Insurance. *Journal of Risk and Insurance*, 85(1): 7-34.
- Kaufman BE, Hotchkiss J (2005). *The Economics of Labor Markets*, 7th rev. ed. Mason, OH: South-Western, Division of Thomson Learning.
- LaMontagne C (2014). Should I Choose A High Or Low Deductible Health Insurance Plan? *Forbes*, December 9. Available from: <https://www.forbes.com/sites/christinalamontagne/2014/12/09/should-i-choose-a-high-or-low-deductible-health-insurance-plan/?sh=629138164ef7>
- Long M, Rae M, Claxton G (2016). A Comparison of the Availability and Cost of Coverage for Workers in Small Firms and Large Firms: Update from the 2015 Employer Health Benefits Survey. Available at: <https://www.kff.org/private-insurance/issue-brief/a-comparison-of-the-availability-and-cost-of-coverage-for-workers-in-small-firms-and-large-firms-update-from-the-2015-employer-health-benefits-survey/>.
- Mayer K (2020). Employer-Sponsored Healthcare Costs Top \$21K a Year. *Human Resource Executive*. Available from: <https://hrexecutive.com/employer-sponsored-healthcare-costs-top-21k-a-year/>
- Mendoza, RL (2017). Information asymmetries and risk management in healthcare markets: The US Affordable Care Act in retrospect. *Journal of Economic Issues*, 51(2): 520–540.
- Mendoza RL (2024). The virtues and vices of association health plans. *Journal of Insurance Regulation*, 43(3): 1–20.
- Mendoza, RL (2025). From financial risk management to promotive health: A paradigm shift in employer-sponsored health insurance? *Journal of Health Management*, 27(3): forthcoming. Available at: <https://journals.sagepub.com/doi/full/10.1177/09720634241245786>
- Miller S (2016). Small Businesses Are Dropping Healthcare coverage; Large Employers Hold Steady. *SHRM Online Benefits*, August 4. Available from: <https://www.shrm.org/resourcesandtools/hr-topics/benefits/pages/small-business-health-coverage.aspx>.
- Morrisey, MA, Kilgore ML, Becker DJ, Smith W, Delzell E (2013). Favorable Selection, Risk Adjustment, and the Medicare Advantage Program. *Health Services Research*, 48(3): 1039–1056.
- Morrisey, MA (2020). *Health Insurance*, 3rd ed. Chicago, Illinois: Health Administration Press.
- National Conference of State Legislatures (NCSL) (2018). Small and Large Business Health Insurance: State & Federal Roles. Available from: <https://search.health/small-business-health-insurance.aspx>
- Office of the President of the United States. Executive Order 13813 (2017). Available at: <https://www.federalregister.gov/documents/2017/10/17/2017-22677/promoting-healthcare-choice-and-competition-across-the-united-states>
- Park CH (2000). Prevalence of Employer Self-Insured Health Benefits: National and State Variation. *Medical Care Research and Review*, 57(3): 340-360.
- Rosen S (1986). The Theory of Equalizing Differences. In O Ashenfelter & R Layard (Eds), *The*

- Handbook of Labor Economics* (pp. 641–692). New York: Elsevier.
- U.S. Census Bureau (2016). *Annual Survey of Entrepreneurs (ASE) - Characteristics of Businesses: 2016 Tables*. Available from: <https://www.census.gov/data/tables/2016/econ/ase/2016-ase-characteristics-of-businesses.html>.
- U.S. Department of Health and Human Services (HHS) (2011). Insurance Standards Bulletin Series—INFORMATION: Application of individual and group market requirements under Title XXVII of the Public Health Service Act when insurance coverage is sold to, or through, associations. Baltimore, MD: HHS. Available at: https://www.cms.gov/CCIIO/Resources/Files/Downloads/association_coverage_9_1_2011.pdf
- U.S. Department of Labor (DOL) (2018). Definition of Employer—Association Health Plans” (2018 AHP Rule), 29 CFR 2510.3–5 (83 FR 28912). Available at: <https://www.govinfo.gov/content/pkg/FR-2018-06-21/pdf/2018-12992.pdf>
- U.S. Internal Revenue Code (IRC), Title 26, United States Code (USC).
- Williamson OE (1979). Transaction-cost economics: The governance of contractual relations. *The Journal of Law & Economics*, 22(2): 233-261.

This page intentionally left blank.

Small Business Administration Lending and Economic Performance

Chu-Sheng Tai*

Abstract

This paper analyzes a panel of state-level data from 1990 to 2023 to determine whether lending by the U.S. Small Business Administration (SBA) through two specific loan programs, 504 and 7(a), significantly impacts state-level real GDP and unemployment rates. Additionally, it seeks to understand how these impacts changed during the global financial crisis (GFC) from 2007 to 2009, which originated from the U.S. subprime mortgage crisis. Using a dynamic panel data model with the Generalized Method of Moments (GMM) estimator, we find that SBA lending positively promotes state-level real GDP and reduces unemployment rates, and these effects are statistically significant and economically important. Specifically, a 1% increase in SBA 7(a) loans per job supported correlates with a decrease of over six basis points in the state unemployment rate. However, this positive effect diminished during the GFC, resulting in a 0.002% decrease in state-level real GDP and a 0.03% increase in the unemployment rate across both loan programs.

Key Words: Small Business Administration, Small Business Lending, Dynamic panel data, GMM

JEL Classifications: G38, H81, L26

I. Introduction

Small businesses are often acclaimed as the engines of both developing and developed economies. There is a widely held belief in their crucial role in economic growth, job creation, boosting foreign trade, and alleviating poverty. Consequently, policymakers frequently highlight small businesses as vital sources of employment growth. However, small business owners continue to face significant challenges in obtaining the necessary funds to execute their innovative ideas and become viable enterprises. In 2022, two out of every three business owners in the U.S. who sought credit did not receive the amount they needed.¹ These persistent gaps in access to capital stem from two issues related to information asymmetry: adverse selection and moral hazard, as identified by Stiglitz and Weiss (1981).² To mitigate the credit constraints faced by small businesses, the U.S. Small Business Administration (SBA) provides federal loan guarantees through its two key

* Department of Accounting and Finance , Jesse H. Jones School of Business, Texas Southern University 3100 Cleburne Avenue, Houston, TX 77004, USA. Tel: +1 713 313 7308; Fax: +1 713 313 7722; e-mail: chu-sheng.tai@tsu.edu

¹ SBA press release, April 12, 2023.

² Stiglitz and Weiss (1981) argued that adverse selection impedes the ability of markets to allocate credit using just the lending rate because it increases the proportion of high-risk borrowers in the pool of prospective borrowers. On the other hand, moral hazard reduces the ability of rates alone to clear lending markets because once the loan is extended, the actions of borrowers is not independent of the lending rate. As a result, difficulties in obtaining adequate information about the parties involved in a transaction may explain the inefficient allocation of small business loans. In such an environment, there are no market forces leading supply to equal demand, and credit is rationed. (Craig et al. 2009).

lending programs, the 7(a) and 504 Loan Programs. These programs offer loan guarantees to eligible small businesses that are unable to secure credit elsewhere. Over the years, the volume of guaranteed loans has steadily increased. In fiscal year 2023, the SBA guaranteed a record \$33.94 billion in loans for small businesses, representing an 11.4% increase compared to the \$30.46 billion guaranteed in fiscal year 2017.

Given the advancements in computer technology and continuous innovations in communication that have significantly improved the efficiency of credit markets—especially for small businesses—we would expect a considerable decrease in information costs stemming from problems of asymmetry in the small business lending market. Moreover, if SBA loan guarantees effectively mitigate credit market frictions, they will lead to improved capital allocation within the economy, which, in turn, should have a positive impact on overall economic performance.

Conversely, if SBA loan guarantees fail to alleviate credit rationing in lending markets, we should not observe a significantly positive correlation between the level of SBA guarantees and economic performance. Therefore, the steady growth of SBA loan guarantee programs raises the question of whether there is any observable impact of SBA activities on U.S. economic growth and employment.

Empirical evidence on this issue remains unclear, warranting further investigation. Some research indicates that SBA-guaranteed lending programs positively affect economic performance, such as per capita income and employment growth at the local market level (Craig et al., 2007, 2008; Cortes, 2010; Orzechowski, 2020). Conversely, other studies argue that SBA-guaranteed lending programs have minimal effects on income or employment improvements in recipient areas (De Rugy, 2007; Lee, 2018; Higgins et al., 2021; Rupasingha et al., 2019). Most of these studies typically examined annual SBA lending at the county or metropolitan statistical area (MSA) levels using data predating 2002 (Craig et al. 2007, 2008; Armstrong et al., 2014; Lee, 2018; Brown & Earle, 2017), with only a few studies investigating the relationship between SBA lending programs and state-level economic performance (Cortes, 2010; Orzechowski, 2020). A recent study by Brown and Earle (2017) explored the impact of access to SBA loans on firm-level employment growth. Based on OLS and instrumental variable regression models, their results suggested that firms receiving SBA loans increased employment by an average of 3.0 to 3.5 employees for every \$1 million in SBA loans received during the first three years after the loans were granted. However, their study was limited to immediate firm-level effects and did not consider the broader impacts of SBA loans on the local economy, including GDP growth and unemployment rates.

The objective of this study is to fill a gap in the literature by providing separate estimates of the impacts of the SBA's 7(a) and 504 loan programs on state-level economic performance, which have often been analyzed together. Additionally, we expand our analysis to cover the period of the 2007-2009 global financial crisis to examine how that crisis affected SBA lending, economic growth, and employment—an area that previous studies have overlooked.

Using a panel of state-level data from 1990 to 2023, we assess whether SBA lending through the two distinct loan programs, 504 and 7(a), significantly impacts state-level GDP and unemployment rates and how this impact changed during the 2007-2009 financial crisis related to the U.S. subprime mortgage crisis. Given the consistent growth of the SBA loan guarantee programs over the years, our empirical results based on the latest data offer important policy implications: Can the ongoing expansion of these government loan guarantee programs be justified?

Based on a dynamic panel data model with a Generalized Method of Moments (GMM) estimator that accounts for potential endogeneity in SBA loan distributions (as noted by Krishnan

et al. (2015) and Brown and Earle (2017)), we find that SBA lending has a statistically significant and economically important positive impact on state-level output and unemployment reduction. Specifically, a 1% increase in SBA 7(a) loans per job supported correlates with a decrease of over six basis points in the state unemployment rate. To verify the robustness of our results, we also considered alternative measures of SBA loans, and the findings remained consistent.

The remainder of the paper is organized as follows. Section II provides an overview of SBA major lending programs. A brief review of the academic literature regarding the impact of SBA lending on the U.S. economic performance is presented in Section III. The empirical model and the econometric methodology are covered in Section IV. Section V discusses the data. Section VI reports and discusses the empirical results. The robustness tests are given in Section VII. Finally, section VIII concludes the paper.

II. The Small Business Administration Loan Programs

The SBA, an independent agency of the United States government, was created on 30 July 1953 by the enactment of Public Law 163. The mission of the SBA is "to maintain and strengthen the nation's economy by enabling the establishment and viability of small businesses and by assisting in the economic recovery of communities after disasters." The SBA's two main business lending programs are the 7(a) guaranteed loan program and the 504 loan program. The 7(a) loan program is the most basic and most significant among the SBA's business loan programs, which provides financial support to small businesses unable to secure financing under reasonable terms from traditional avenues. Loans from the 7(a) program are only available on a guaranty basis, and the program guarantees up to 85% of loans with a maximum guaranty of up to \$150,000 and 75% of loans greater than \$150,000 made by partnering lenders, which may include banks, credit unions, and other financial institutions. Under the guaranty concept, commercial lenders make and administer the loans, and small businesses apply to lenders for their financing. The lender decides whether it will make a loan internally or if the application has some weaknesses which, in the lender's opinion, mean the loan will require an SBA guaranty before it will be underwritten. The guaranty that the SBA provides is available only to the lender. It assures the lender that in the event of a payment default, the government will reimburse the lender for its loss, up to the percentage of SBA's guaranty. Under the 7(a) program, the borrower remains obligated for the full amount due. This federal guarantee reduces the risk to the lender, thereby encouraging them to provide loans to small businesses that might otherwise be considered too risky. The loan funds can be used for a variety of purposes, including working capital (both short and long term), refinancing debt, and purchasing furniture, fixtures, and supplies. This flexibility reflects the SBA's recognition of the diverse needs of small businesses and the barriers they face in accessing traditional credit markets. However, some businesses that are ineligible for this program, such as real estate investment firms (where property is held for investment purposes), dealers of rare coins and stamps, and lending institutions like banks.

The 504 Loan Program is a direct debentures program that provides fixed-rate loans to small businesses. This loan program is specifically designed for the long-term financing of fixed assets like real estate or heavy machinery and equipment but provides only 40% of the funding. Fifty percent of the remaining funding usually comes from banks and 10% from the small business owner depending on the SBA rules and regulations. SBA 504 financing is provided through a certified development company (CDC), a nonprofit corporation set up to contribute to the

economic development of its community. CDCs work with the SBA and private-sector lenders to provide financing to small businesses. The 504 loans cannot be used for working capital or inventory, underscoring the program's explicit focus on job creation and retention via physical capital expansion. For instance, borrowers are generally required to create or retain one job for every \$65,000 guaranteed by the SBA. Understanding the 7(a) and 504 loan programs in this detailed context helps shed light on the mechanisms through which SBA loans may impact economic and employment outcomes and the potential reasons behind the variations in their effectiveness. It also underscores the significant role the SBA plays in addressing the capital access issues faced by small businesses.

III. Literature Review

Previous studies on the impact of SBA lending on economic performance primarily focused on data up to 2013. For instance, Hicks (2004) used data from the Community Reinvestment Act (CRA) to analyze the effect of CRA-reported loans on employment in 55 counties of West Virginia during the 1996-98 period. His growth model regressed county employment on CRA-reported loans of less than \$100,000, human capital, public capital, county distress rankings, a trend variable, and a spatial autocorrelation adjustment. He found that loans to small businesses had a positive and statistically significant effect, but only for firms with five to nine employees. Craig et al. (2007) examined SBA loan activity in metropolitan statistical areas (MSAs) and non-MSA locations using annual data from 1991 to 2001. They identified a positive relationship between per capita personal income growth and SBA loans using a GMM regression model, noting that the effects were somewhat stronger in non-MSA locations compared to MSAs. Although statistically significant, the economic impact was minor based on the estimated coefficients. In a subsequent study (Craig et al., 2008), they analyzed the relationship between SBA loans and employment in MSA locations, again using annual data from 1991 to 2001 and applying a fixed-effects ordinary least squares (OLS) model. After controlling for the density of financial institutions, local economic structure, and the retail industrial share of the local economy, they found a statistically significant and positive impact of SBA loans on local employment, especially in lower-income MSAs compared to higher-income ones. Cortes (2010) studied the impact of SBA loans on various small business activity indicators, such as employment rates and per capita income, while controlling for other determinants of state economic growth. Using U.S. state-level data from 1986 to 2008 and estimating three different dependent variables—income growth, small firm growth, and employee growth—his regression results indicated that the impact of SBA lending on per capita income was statistically insignificant. Conversely, SBA loans had a positive and significant impact on the growth of small businesses and the number of employees within those firms. Notably, his study did not account for potential endogeneity in the model nor differentiate between the impacts of the two distinct SBA lending programs on state economic performance. Armstrong et al. (2014) analyzed the role of SBA lending in relation to employment and financial market development using a generalized least squares (GLS) fixed-effects regression model with annual data from 1991 to 2001. They found that SBA lending had a statistically significant positive effect in areas with relatively underdeveloped financial markets, concluding that the SBA guarantee helped improve credit access in these less developed regions. Brown and Earle (2017) estimated the firm-level impact of access to SBA loans on employment growth at the county level. Using OLS and instrumental variable regression models with annual data from 1991 to 2009, their results indicated that firms that received SBA loans experienced growth in employment. Lastly,

Orzechowski (2020) examined the relationship between SBA lending programs and state-level employment using a dynamic GLS panel model. His analysis used quarterly state-level data from the U.S. covering the years 1990 to 2013. The results indicated a statistically significant positive relationship between growth in SBA lending per capita and changes in state civilian employment rates. However, while these relationships were statistically significant, they were not economically meaningful. An analysis comparing states with high and low personal incomes revealed no significant differences in the relationship between SBA lending and employment across these groups. It is important to note that this study focused solely on the connection between SBA lending and state-level employment, without considering output growth.

IV. Empirical Model and Methodology

To examine whether the SBA lending has a statistically significant relationship with state-level economic growth and employment, we estimate the following equation:

$$Y_{i,t} = \beta_0 + \beta_1 Y_{i,t-1} + \beta_2 SBA_{t-1} + \beta_3 (SUB * SBA)_{t-1} + \beta_4 DEP_{i,t-1} + \beta_5 LQ_{i,t-1} + \beta_6 HPI_{i,t-1} + \varepsilon_{i,t} \quad (1)$$

Where $Y_{i,t}$ represents the real state-level gross product per capita (*RSGP*) or unemployment rate (*UNEMP*) of state i at time t . The primary variable of interest on the right side of equation (1) is, SBA_{t-1} , which is the lagged SBA loans, measured as the loan amount per job supported. In addition, the variable SBA loans is categorized into two groups—504 loans and 7(a) loans—to determine whether the effects of SBA loans on the real state-level GDP and unemployment rate vary based on the type of SBA programs. Therefore, the first two hypotheses of this study are as follows:

H1: SBA lending will increase the real state-level gross product per capita (i.e., $\beta_2 > 0$)

H2: SBA lending will reduce the state-level unemployment rate (i.e., $\beta_2 < 0$)

Another variable of interest is $(SUB * SBA)_{t-1}$ where SUB is a dummy variable equal to one during the 2007-2009 global financial crisis (GFC) originated from the U.S. subprime mortgage crisis and zero otherwise. A statistically significant negative β_3 coefficient indicates that the contribution of SBA lending to the state-level economic growth diminished during the crisis. Therefore, the next two hypotheses of this study are as follows:

H3: The GFC has a negative effect on the contribution of SBA lending to the state-level gross product per capita (i.e., $\beta_3 < 0$)

H4: The GFC has a negative effect on the contribution of SBA lending in reducing state-level unemployment rate (i.e., $\beta_3 > 0$)

To controls for national economic conditions and local market structure, we include three additional regressors. First, $DEP_{i,t-1}$ is bank deposit per capita which is a proxy for the financial development and competitiveness of banking sector in the local market (Craig et al. 2007; Corets 2010). Second, $LQ_{i,t-1}$, is a variable of location quotient for manufacturers calculated as a ratio of the share of manufacturing employment in state employment to the share of overall manufacturing employment in U.S. It is intended to measure how concentrated a particular industry was in a state

compared to the national average because local industrial composition is an important determinant of local employment (Cortes, 2010; Ramsey, 2018). A ratio greater than one indicates that a state's manufacturing sector accounts for a larger share of state employment as compared to that of the nation. Third, $HPI_{i,t-1}$ is the house price index which can be used to measure how the economy will likely to perform. A higher price index generally signals that more jobs will be created which stimulates consumer confidence and spending. This paves the way for greater aggregate demand, boosting GDP and overall economic growth. We take the log of all the variables defined in equation (1) in the estimation.

The dynamic nature with the lagged dependent variable included as a regressor in equation (1) suggests the possibility of an endogenous bias, as outlined by Nickell (1981). When dependent variable exhibits high persistence (i.e., the persistence coefficient β_1 is positive and statistically significant), using static models such as pooled OLS and fixed effect produce biased and inconsistent results (Blundell and Bond, 1999). To address potential endogeneity issues, Arellano and Bond (1991) and Blundell and Bond (1998) developed the generalized method of moments (GMM) model, which can be used for dynamic panel data. The GMM model provides consistent results in the presence of different sources of endogeneity, namely “unobserved heterogeneity, simultaneity and dynamic endogeneity” (Wintoki, Linck, & Netter, 2012). The GMM model removes endogeneity by internally transforming the data, and two kinds of transformation methods are available. The first method, known as first-difference transformation, where a variable's past value is deducted from its current value, could result in the loss of too many observations and an increase in the gap between observations when there are missing data (Roodman, 2009). The second method, known as second-order transformation, applies “forward orthogonal deviations” which means that instead of subtracting the previous observations of a variable from its current value, the second-order transformation subtracts the average of all future available observations of a particular variable (Roodman, 2009). Using the second-order transformation, researchers can prevent unnecessary data loss. Therefore, in the case of an unbalanced panel dataset, a GMM model with the second-order transformation provides more efficient and consistent estimates for the involved coefficients and is adopted in this study.³

V. Data

To examine the impact of SBA lending on state-level employment and economic growth, we construct a panel of U.S. data that includes observations on 51 states from 1990 to 2023. Our dependent and independent variables are constructed as annual averages over this sample period for a total of 1734 observations. We utilize the data from different sources. The first source of data on loan-specific data for all SBA-approved 7(a) and 504 loans was obtained directly from the SBA and consists of 211,943 504 loans and 1,764,630 7(a) loans for a total of 1,976,573 loans approved over the sample period. SBA loans were grouped at the state-level and across time in order to add to the SBA literature by exploring state-level economic markets that have not been fully studied by the existing SBA literature. Table 1 reports the data collected on these two SBA lending programs by calculating the total loan volume, average loan amount, and the number of

³ The simulation results show that the GMM estimator of the model transformed by the forward orthogonal deviation tends to work better than that transformed by the first difference (K. Hayakawa, 2009). Therefore, we estimate the GMM estimator of our dynamic panel data model transformed by both first difference and forward orthogonal deviation and we only select the model transformed by the first difference if the model transformed by the forward orthogonal deviation is rejected.

jobs supported per loan separately under each program for each state and Table 2 reports the annual averages of those loan data across all 51 states over the 1990-2023 sample period.

Table 1: SBA 504 and 7(a) Loans by U.S. States

The table shows the total quantity of the loans approved, the average loan amount per loan, and the average numbers of job supported per loan across 51 U.S. states during 1990-2023. States are given by their U.S. postal abbreviation.

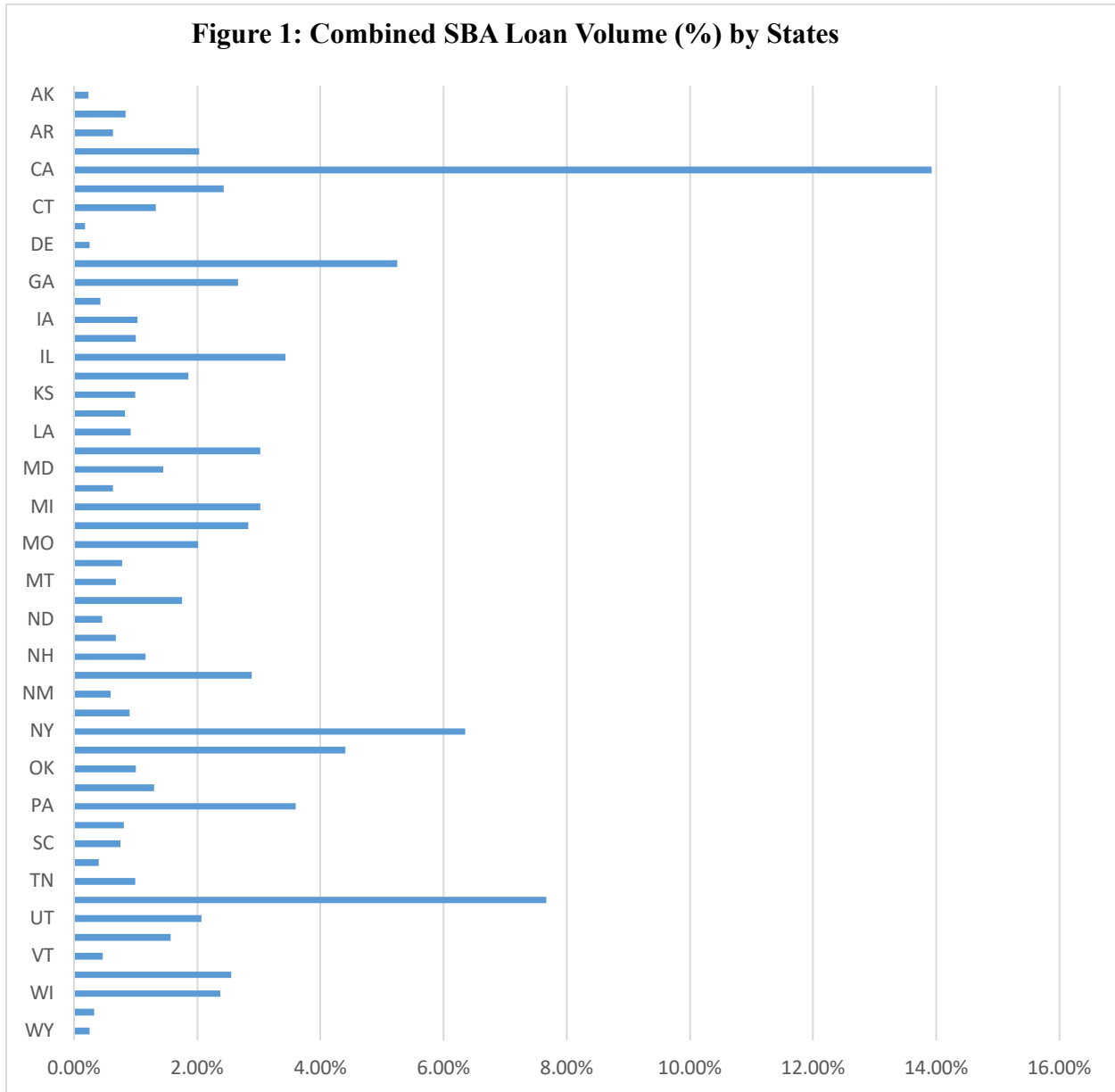
State	SBA 504 Loans			SBA 7(a) Loans			
	Qty	Gross Approved	Jobs Supported	Qty	Gross Approved	Guaranteed	Jobs Supported
<i>AK</i>	468	538228.7	13	4084	354821.6	270617.7	9.35
<i>AL</i>	2989	534709.5	15	13615	352074.4	268761.5	8.91
<i>AR</i>	778	635648.5	21.27	11740	338344.9	254875.8	7.18
<i>AZ</i>	4385	592327.8	19.82	35721	377410.7	282138	12.5
<i>CA</i>	41848	708816.7	20.72	233367	395028.1	298393.3	10.74
<i>CO</i>	6427	557208.3	17.95	41564	345726.9	260483	9.78
<i>CT</i>	1830	493991.7	16.79	24358	258805.6	188515.1	8.86
<i>DC</i>	274	753748.8	17.11	3286	279485.4	209303.7	10.55
<i>DE</i>	218	646794	19.75	4815	243178	180054.5	8.56
<i>FL</i>	14934	546931.4	12.92	88889	343153.7	258374.7	9.71
<i>GA</i>	5715	617115.2	15.97	46970	483097.5	366644.7	10.03
<i>HI</i>	1077	537834	13.25	7474	136899.4	99721.03	6.88
<i>IA</i>	2625	478898.3	16.47	17843	246210	186904.8	8.52
<i>ID</i>	3083	406352.4	13.88	16825	230202.4	174009.3	7.86
<i>IL</i>	9043	592481.5	17.53	58765	299453.4	223740.5	9.12
<i>IN</i>	4828	452671.1	18.72	31811	269194.9	199463.7	8.96
<i>KS</i>	1764	493470.8	17	17879	260604.7	195893.4	8.97
<i>KY</i>	1162	532711.9	28.27	15249	253444.9	188323.2	8.71
<i>LA</i>	1226	531508.3	18.56	17025	326453.6	247338.7	9.69
<i>MA</i>	6158	507062.9	16.69	53606	166739.1	118700.7	7.85
<i>MD</i>	1904	705772.6	21.07	26660	251427	186985.1	9.73
<i>ME</i>	1455	408980.7	12.58	11022	169688.1	123064.5	7.57
<i>MI</i>	5163	519857	22.21	54544	270965	198317.3	8.31
<i>MN</i>	8632	490933.6	19.16	47236	254291.2	189195	10.23
<i>MO</i>	4395	497370.9	18.15	35435	265762.9	201480.4	9.16
<i>MS</i>	551	840880.6	24.5	14906	287068	218820.7	6.49
<i>MT</i>	881	501860.6	14.74	12514	234843.4	177806	7.01
<i>NC</i>	3680	561628.3	12.88	30954	358666.6	269445.6	9.67
<i>ND</i>	1532	484627.2	15.28	7504	210573.5	151720.8	7.99
<i>NE</i>	1327	544299.3	19.87	12005	232926.1	174851.3	8.72
<i>NH</i>	3715	355598	14.87	19263	154110.9	110508.9	7.46

<i>NJ</i>	2206	822131.1	18.46	54818	323786.8	242261.1	9.07
<i>NM</i>	1367	495620.6	16.21	10380	303685.6	229597.3	9.6
<i>NV</i>	3489	574510.4	18.98	14293	306473.8	230786.8	10.6
<i>NY</i>	8032	636874.8	20.91	117496	220229.7	160884.8	7.78
<i>OH</i>	7045	448173.6	15.46	80055	208911.5	152208.4	7.63
<i>OK</i>	1659	523953.6	21.75	18111	307723	230030.1	8.98
<i>OR</i>	2308	636676.8	15.23	23368	295982.1	222540.1	8.53
<i>PA</i>	3218	631187.6	18.72	67992	252110.9	187080.9	8.79
<i>RI</i>	1213	459831.1	18.59	14709	194493.6	143797.5	8.82
<i>SC</i>	1398	650226.3	19.34	13432	392575.9	297700.6	9.58
<i>SD</i>	1591	468148.8	15.78	6385	232280.7	177357.5	6.85
<i>TN</i>	1393	696263.4	19.59	18257	368323.8	278818.4	9.5
<i>TX</i>	9904	717173.2	23.49	141577	379268.1	285432.4	9.94
<i>UT</i>	7449	505872.9	16.21	33368	260251.4	194157.8	10.84
<i>VA</i>	4348	534566.3	12.42	26658	312308.3	234331.7	9.92
<i>VT</i>	578	454887.6	18.42	8610	162442	117428.7	7.37
<i>WA</i>	4516	650817.1	13.31	45942	335412.9	252067	8.62
<i>WI</i>	5397	605105.6	16.25	41518	284794.7	212837.3	9.16
<i>WV</i>	211	562352.5	19.11	6294	218715.7	160676.4	7.02
<i>WY</i>	554	561592.1	16.37	4438	274743.6	207784.6	8.42

As can be seen from Figure 1, most of the SBA lending is concentrated in California, which has 13.92% of the total SBA volume. Texas (7.66%), New York (6.35%), and Florida (5.25%) follow California with relatively large shares when compared to the remaining states. Figure 2 plots the time series of SBA lending over the sample period and as can be seen, the volume of 7(a) loans is always higher than that of 504 loans and the total TBA lending reached the highest point just right before the start of 2007-2009 US sub-prime mortgage crisis in 2006. Figure 3 depicts the average loan amount approved for each program, and it shows that although the number of loan approved is always higher for 7(a) loan program, but the average loan amount approved for each program is always higher for 504 loan program which is not surprising given the nature of 504 loan which focuses more on physical capital investments instead of inventory purchases under 7(a) loan program. Figure 4 exhibits the average number of jobs supported per loan approved under each loan program over time, and this average is always higher for the 504 loan program but it was gradually decreasing up to 2009 and afterward the average number of jobs supported is very close for both loan programs.

The data to measure economic conditions including the real gross state product per capita is directly obtained from the Bureau of Economic Analysis (BEA). The unemployment rate and employment data used to calculate the location quotient for manufacturing are from the Bureau of Labor Statistics (BLS). The house price index, one of many economic indicators that investors use to keep a pulse on broader economic trends, is downloaded from Federal Housing Finance Agency (FHFA) website. The data used to measure the financial development in each state, bank deposit per capita, is obtained from the Federal Deposit Insurance Corporation's (FDIC) annual summary of deposit data files. The definitions and the descriptive statistics of the variables used in the estimation is in Table 3, and Table 4 shows the correlations among the log values of independent variables used in the estimation. Most of the correlations are very low except the

correlation between the (log) SBA 504 gross loan per job supported and the (log) house price index at 0.6198. We also calculate the Variance Inflation Factor (VIF) for each independent variable and the VIFs are all less than 5 and range from 1.10 for (log) bank deposit per capita to 3.04 for the (log) house price index, suggesting no presence of multicollinearity among our dependent variables.⁴



⁴ The results of multicollinearity tests not reported here but are available upon request.

Figure 2: Loan Volume Approved by Years

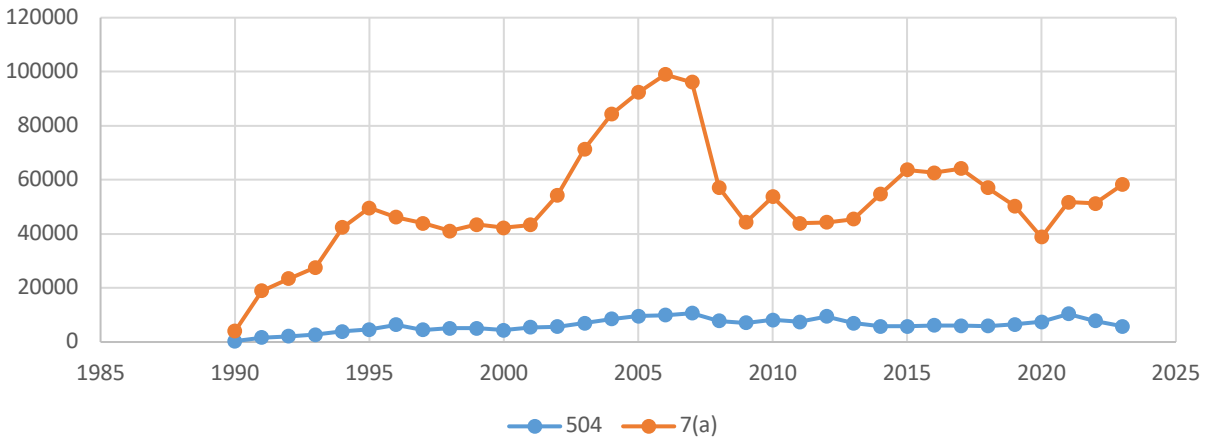


Figure 3: Average Loan Amount by Years

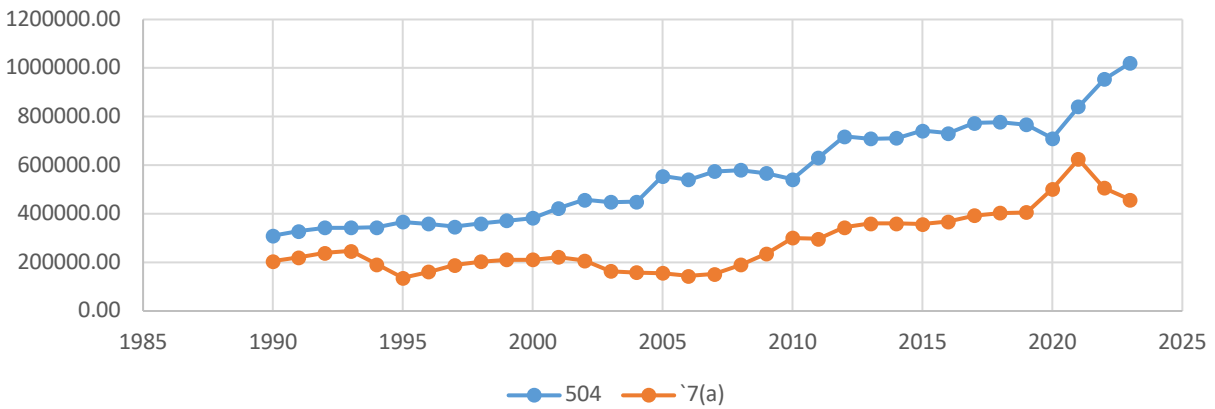


Figure 4: Number of Jobs Supported per Loan

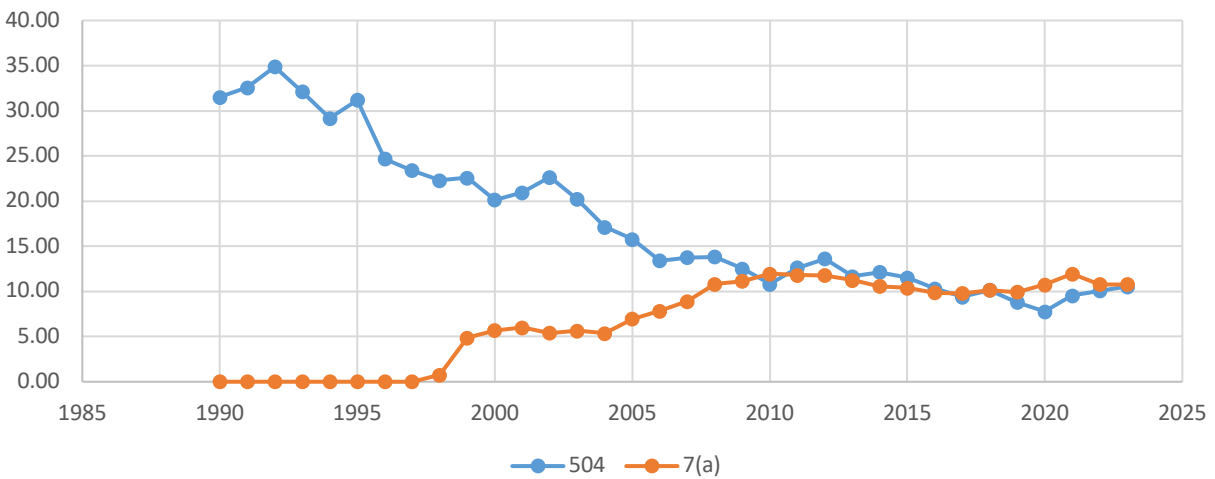


Table 2: SBA 504 and 7(a) Loans by Years

The table shows the total quantity of the loans approved, the average loan amount per loan, and the average numbers of job supported per loan by years during 1990-2023.

Year	SBA 504 loan Program			SBA 7(a) Loan Program			
	Qty	Gross Approved	Jobs Supported	Qty	Gross Approved	Guaranteed	Jobs Supported
1990	329	310118.9	31.52	3995	204255.5	165131.5	NA
1990	329	310118.9	31.52	3995	204255.5	165131.5	NA
1991	1631	329081.1	32.57	18909	220240	178484.3	NA
1992	2095	342914.7	34.87	23378	239259.3	196003.3	NA
1993	2764	343004.3	32.12	27585	246912.3	196545.4	NA
1994	3898	344758.4	29.18	42384	190968.9	147611.3	NA
1995	4637	367132.7	31.2	49538	136640.1	107868.1	NA
1996	6340	358662.4	24.67	46131	161401.3	119026.1	NA
1997	4434	346977	23.39	43952	188840.4	137318.2	0.01
1998	5064	359799.8	22.31	41106	203069.6	149082.2	0.77
1999	5132	372472.9	22.6	43392	211430.7	150316.7	4.87
2000	4423	381873	20.13	42262	211478.4	151922.9	5.67
2001	5420	422763.2	20.94	43269	221980.3	163860.1	6
2002	5714	457760.9	22.66	54299	207087.2	152812.8	5.4
2003	6953	447725.5	20.22	71374	164113.2	118937.2	5.63
2004	8497	449925.9	17.12	84325	158747.8	112971.6	5.38
2005	9572	555257.8	15.78	92320	155456.4	110138.5	6.94
2006	9887	540970.5	13.4	98950	143579.1	100205.5	7.83
2007	10647	574800.8	13.77	96115	151403.8	106122.2	8.91
2008	7797	579531.1	13.83	57196	190387	136047.9	10.81
2009	7149	566765.8	12.51	44242	236106.7	192662.2	11.14
2010	8105	541397.8	10.82	53784	300748.5	246996.6	11.94
2011	7437	630877.3	12.63	43906	296791.4	213098.7	11.81
2012	9516	717733.4	13.63	44254	344728.7	251059.9	11.77
2013	6974	709435.4	11.67	45486	360298	264191.7	11.23
2014	5805	712214	12.14	54714	359451.1	263457.3	10.58
2015	5755	742118	11.51	63658	357509.7	262174.3	10.39
2016	6140	732201.7	10.31	62657	367734.6	269394.8	9.85
2017	6084	773175.7	9.4	64107	393046.3	288210.1	9.78
2018	5886	777457.2	10.16	57009	403055.5	295599.9	10.13
2019	6496	767412.4	8.81	50311	405749.6	297033.1	9.92
2020	7399	710142.7	7.79	38836	502896.8	366659	10.76
2021	10441	840441.1	9.55	51720	624924	518744.5	11.94
2022	7777	953688.7	10.08	51240	505868	368333	10.78
2023	5745	1021506	10.52	58226	457283.6	332937.3	10.77

Table 3: Descriptive Statistics

Variable	Source	Mean	Std. dev.	Max	Min
Real State GDP	BEA	51279.11	3640.1	57073.76	43502.98
Unemployment rate	BLS	5.36	1.41	8.68	3.29
504 gross loan per job supported	SBA	47440.71	32076.96	107359.5	11559.54
7(a) gross loan per job supported	SBA	188767.7	646385.7	3327567	17659.51
7(a) guaranteed loan per job supported	SBA	136801.1	464411.2	2388046	12374.42
Bank deposit per capita	FDIC	48.76	38.93	121.89	10.38
Location quotient for manufacturing	BLS	0.95	0.03	0.99	0.91
House price index	FHFA	181.43	58.32	335.86	100

BEA: Bureau of Economic Analysis; BLS: Bureau of Labor Statistics; FHFA: Federal Housing Finance Agency; FDIC: Federal Deposit Insurance Corporation. The data for real state GDP per capita obtained from BEA directly is only available from 1997 to 2019. The date for the number of jobs supported under 7(a) loan program is not available until 1997, so the descriptive statistics calculated for 7(a) gross loan per job supported and 7(a) guaranteed loan per job supported are for the 1997-2023 sample period.

Table 4: Correlations

Panel A	$\ln(SBA)^a$	$\ln(DEP)$	$\ln(LQ)$	$\ln(HPI)$
$\ln(SBA)^a$	1.0000			
$\ln(DEP)$	0.0821***	1.0000		
$\ln(LQ)$	-0.0497	0.2609***	1.0000	
$\ln(HPI)$	0.6198**	0.0132	-0.2293***	1.0000
Panel B	$\ln(SBA)^b$	$\ln(DEP)$	$\ln(LQ)$	$\ln(HPI)$
$\ln(SBA)^b$	1.0000			
$\ln(DEP)$	-0.0567	1.0000		
$\ln(LQ)$	0.0142	0.2609***	1.0000	
$\ln(HPI)$	-0.3431***	0.0132	-0.2293***	1.0000

$\ln(SBA)^a$ is the log of SBA 504 gross loan per job supported; $\ln(SBA)^b$ is the log of SBA 7(a) guaranteed loan per job supported; $\ln(DEP)$ is the log of bank deposit per capita, $\ln(LQ)$ is the log of the location quotient; $\ln(HPI)$ is the log of house price index. ** = significance at 5% level, and *** = significance at the 1% level. The data for real state GDP per capita obtained from BEA directly is only available from 1997 to 2019, so the correlations calculated are for the 1997-2019 sample period.

VI. Empirical results

Table 5 presents the results of the impact of SBA lending on the state-level real GDP and unemployment rate for 504 and 7(a) loan programs separately in Panels A and B, respectively. Before we investigate the statistical significance of the parameter estimates specified in equation (1), we need to check if our dynamic panel data model with GMM estimator is appropriate by testing the overidentifying restriction based on Hansen's J-test. The p-values of this test ranging from 15.16% to 29.68% reported in Table 5 imply that the instruments used in our GMM estimation are valid. We can now examine the statistical significance of the parameter estimates in the model. Notice from the table that the coefficients (β_1) on the lagged dependent variable, $\log(RSGP_{t-1})$, for 504 loans (0.8316) and 7(a) loans (0.9054) are positive and statistically significant at the 1% level. Similarly, coefficients (β_1) on $\log(UNEMP_{t-1})$ for 504 loans (0.6843) and 7(a) loans (0.5933) are also positive and statistically significant at 1% level. These results imply the persistence of both dependent variables and thus justify the use of the dynamic panel data model. The main interest of this study is the β_2 coefficient on $\log(SBA_{t-1})$ which measures the impact of the SBA lending calculated as the (log) amount of SBA approved loan per job supported on the (log) stat-level real GDP per capita, $\log(RSGP_t)$, and it is positive and statistically significant at the 1% level for both programs (0.006 for 504 loans and 0.0046 for 7(a) loans), suggesting that the SBA lending does have a positive impact on the economic performance at state level.

In terms of the unemployment rate, the β_2 coefficient on $\log(SBA_{t-1})$ is negative and statistically significant at the 1% level for 7(a) loans (-0.0646). This suggests that SBA lending under the 7(a) loan program does reduce the state-level unemployment rate. Specifically, a 1% increase in the amount of SBA 7(a) loans per job supported is associated with a 0.0646% decrease in the state-level unemployment rate. Conversely, the coefficient β_2 for 504 loans is positive and statistically significant at the 5% level (0.0182), indicating that SBA lending under the 504 loan program may actually increase the unemployment rate. This contradictory finding raises questions about whether SBA lending effectively reduces unemployment. The relationship is complex. Easier access to capital through SBA lending may enable small businesses to expand and hire more workers, representing a positive scale effect. However, it could also lead to capital-labor substitution, which would reduce employment if capital and labor are considered gross substitutes. Furthermore, even if the scale effect is dominant, any increase in employment could be limited if SBA lending crowds out other sources of capital. The overall employment effect may also be reduced if there are negative spillovers onto competing firms, leading to a crowding-out effect. Consequently, it is plausible that the 504 loan program could negatively impact employment since it primarily focuses on fixed assets.

Another interest of this study is the impact of SBA lending during the GFC. As can be seen, the β_3 coefficient on $\log(SUB * SBA)_{t-1}$ is negative and statistically significant at the 1% level for both 504 loans (-0.0022) and 7(a) loans (-0.0019) with respect to the state-level real GDP per capita. With respect to the state-level unemployment rate, β_3 coefficient is positive and statistically significant at the 1% level for both 504 loans (0.0344) and 7(a) loans (0.0289), indicating that the positive impact of both SBA lending programs on the state-level real GDP and employment before the crisis was reduced during the crisis. Because the research on how the SBA lending contributed to the state-level economic performance during the GFC has not been fully exploited in the literature, the empirical results from this study have shed light on this issue and have very significant public policy implications for how SBA guaranteed lending should be

reevaluated. Particularly researchers need to empirically reassess the success of key policy initiatives to mitigate the damage done by the credit crunch associated with the GFC.

Table 5: Dynamic Panel Data Model with GMM Estimator: SBA Gross Loan per Job Supported

This table reports the parameter estimates of the dynamic panel data model specified in equation (1) utilizing the GMM approach for two SBA lending programs: 504 loans in Panel A and 7(a) loans in Panel B.

Panel A: 504 Gross Loan				
	$Y_t = \ln(RSGP_t)$		$Y_t = \ln(UNEMP_t)$	
$\ln(RSGP_{t-1})$	0.8316	(0.0057)***	0.6483	(0.0077) ***
$\ln((SBA_{t-1}))$	0.0060	(0.0012) ***	0.0182	(0.0076) **
$\ln(SUB * SBA_{t-1})$	-0.0022	(0.0001) ***	0.0344	(0.0070) ***
$\ln(DEP_{t-1})$	-0.0025	(0.0007) ***	0.0014	(0.0035)
$\ln(LQ_{t-1})$	0.0671	(0.0057) ***	-0.0196	(0.0161)
$\ln(HPI_{t-1})$	0.0203	(0.0018) ***	-0.1421	(0.0205) ***
HJ-Stat	54.8322		50.3969	
P-Value (J)	0.2019		0.2685	
# of observations	1068		1566	
Panel B: 7(a) Gross Loan				
	$Y_t = \ln(RSGP_t)$		$Y_t = \ln(UNEMP_t)$	
$\ln(RSGP_{t-1})$	0.9054	(0.0058) ***	0.5933	(0.0080) ***
$\ln((SBA_{t-1}))$	0.0046	(0.0007) ***	-0.0646	(0.0023) ***
$\ln(SUB * SBA_{t-1})$	-0.0019	(0.0001) ***	0.0289	(0.0036) ***
$\ln(DEP_{t-1})$	-0.0022	(0.0011)	-0.0076	(0.0026) ***
$\ln(LQ_{t-1})$	0.0301	(0.0064) ***	-0.1500	(0.0263) ***
$\ln(HPI_{t-1})$	0.0105	(0.0022) ***	-0.1781	(0.0136) ***
J-Stat	58.0573		50.6022	
P-Value (J)	0.1516		0.2968	
# of observations	976		1180	

$$Y_{i,t} = \beta_0 + \beta_1 Y_{i,t-1} + \beta_2 SBA_{t-1} + \beta_3 (SUB * SBA)_{t-1} + \beta_4 DEP_{i,t-1} + \beta_5 LQ_{i,t-1} + \beta_6 HPI_{i,t-1} + \varepsilon_{i,t} \quad (1)$$

Where $Y_{i,t}$ represents the real state-level gross product ($RSGP$) per capita or unemployment rate ($UNEMP$) of state i at time t ; SBA_{t-1} , which is the lagged SBA gross loan per job supported; SUB is a dummy variable equal to one during the 2007-2009 global financial crisis (GFC) and zero otherwise. $DEP_{i,t-1}$ is bank deposit per capita, $LQ_{i,t-1}$, is the location quotient; $HPI_{i,t-1}$ is the house price index; $\varepsilon_{i,t}$ represents a mean-zero disturbance term. All the variables in equation (1) have been convert to log values. HJ-Stat is Hansen's J statistic. P-Value (J) is the associated P-value of HJ-test. Robust standard errors clustered at the state level are shown in parentheses ** = significance at 5% level, and *** = significance at the 1% level. The data for real state GDP per capita obtained directly from BEA is only available from 1997 to 2019, so equation (1) was estimated over this sample period when the dependent variable is $\ln(RSGP_t)$.

We next consider our control variables. The β_4 coefficient on $\log(DEP_{i,t-1})$ for the state-level real GDP per capita is statistically significant at the 1% level only for SBA 504 loans with a negative sign (-0.0025), suggesting the negative correlation between the state-level real GDP and

financial development or market liquidity in the state. If the bank deposit per capital can be a proxy for the financial development or liquidity of state economy, this negative correlation seems to be counterintuitive given that a liquid financial market should promote economic growth. However, while bank deposits are essential for the functioning of the banking system, high deposit rates can sometimes lead to lower GDP growth, as they can reduce the amount of money available for lending and investment. Our result is very close to the finding by Driscoll (2004) who used a panel of annual data on the U.S. states and concluded that that shocks to the supply of bank loans generally have a small and often negative effect on state personal income. For the state-level unemployment rate, this β_4 coefficient on $\log(DEP_{i,t-1})$ is significant for SBA 7(a) loans with a negative sign (-0.0076), implying that a well-developed financial market with more liquidity will reduce the state-level unemployment rate, which is one would expect and is also consistent with previous studies including Armstrong et al. (2014), Shaffer (2006), Levine (1997), and Craig et al. (2008).

The coefficients on the remaining two control variables are consistent with our expectations and previous literature. For example, the β_5 coefficient on $\log(LQ_{i,t-1})$ is statistically significant at the 1% level with a positive sign with respect to the state-level real GDP for SBA 504 loans (0.0671) and 7(a) loans (0.0301), implying a 1% increase in the location quotient for manufacturing sector is associated with a 0.0671% (0.0301%) increase in the state-level real GDP under each lending program, respectively. This same coefficient with the expected negative sign is insignificant for the state-level unemployment rate for SBA 504 loans (-0.0196) but statistically significant at the 1% level for 7(a) loans (-0.15), implying that a 1% increase in the state manufacturing employment relative to that of the U.S. is associated with a 0.15% decrease in state-level unemployment rate. Finally, the β_6 coefficient on $\log(HPI_{i,t-1})$ with respect to the state-level real GDP has the expected positive sign and statistically significant at the 1% level for both lending programs (0.0203 vs. 0.0105), and with respect to the state-level unemployment rate, the coefficient also has the correct negative sign and statistically significantly at the 1% level for both lending programs (-0.1421 vs. -0.1781). Based on the statistical significance and the size of the estimated β_6 coefficient on $\log(HPI_{i,t-1})$, it shows that the house price index is not only statistically significant but economically important in predicting the state economic performance.

VII. Robustness check

The results reported in Panel B of Table 5 show a statistically significant and positive impact of SBA lending on both state-level real GDP and employment rate when using the total *gross* loan approved per job supported under SBA 7(a) loan program. Since SBA also provides the information on the portion of gross loans which are guaranteed under 7(a) program, it would be interesting to see if the positive impact found in Panel B of Table 5 may have been over-stated if the guaranteed portion of 7(a) gross loan per job supported is used in the GMM estimation.⁵ The results reported in Table 6 show that the model is not rejected by the Hansen's J-test and in addition the statistical significance and the sign of the coefficient estimates are virtually unchanged. For example, the β_1 coefficient on the lagged dependent variable is statistically significant and positive for both $\ln(RSGP_{t-1})$ and $\ln(UNEMP_{t-1})$. The main interest of β_2 coefficient on $\ln(SBA_{t-1})$ is still statistically significant and positive for state-level real GDP (0.0062) and negative (-0.0607)

⁵ SBA does not provide the information on the guaranteed portion of gross loans under 504 loan program, so we only re-estimate the model for SBA 7(a) loan program for the robustness check.

for the unemployment rate, suggesting that even though the average of SBA 7(a) guaranteed loan per job supported (\$136801.1) is smaller than the average of the overall gross loan per job supported (\$188767.7) shown in Table 3, the positive impact of SBA lending under 7(a) loan program on state-level real GDP and employment remain strong especially in reducing unemployment rate (a 0.0607% drop with a 1% increase in the SBA guaranteed lending per job supported). Regarding the β_3 coefficient on the SBA lending during the GFC, similar to the finding in Table 5, it is statistically significant and negative with respect to state-level real GDP and positive for state-level unemployment rate, implying that the positive impact on promoting state economic growth and reducing unemployment rate from SBA lending programs diminished during the crisis. The β_4 coefficient on $\ln(DEP_{t-1})$ remains statistically insignificant with respect to state-level real GDP, but is statistically significant and negative (-0.0082) with respect to state unemployment rate. The β_5 coefficient on $\ln(LQ_{t-1})$ is statistically significant and positive for both state-level real GDP (0.0254) and unemployment rate (-0.1586), suggesting that a higher concentration of manufacturing industry not only improves state-level real GDP, but also reduces the state unemployment rate. Finally, the β_6 coefficient on $\ln(HPI_{t-1})$ is also statistically significant and positive for both state-level real GDP (0.0129) and unemployment rate (-0.1699), indicating that an increase in the house price index would stimulate state-level economic growth and improve its employment.

Table 6: Dynamic Panel Data Model with GMM Estimator: SBA 7(a) Guaranteed Loan per Job Supported

This table reports the parameter estimates of the dynamic panel data model specified in equation (1) utilizing the GMM approach for SBA 7(a) loans.

	7(a) Guaranteed Loan			
	$Y_t = \ln(RSGP_t)$		$Y_t = \ln(UNEMP_t)$	
$\ln(RSGP_{t-1})$	0.9104	(0.0092) ***	0.5986	(0.0080) ***
$\ln((SBA_{t-1}))$	0.0062	(0.0006) ***	-0.0607	(0.0025) ***
$\ln(SUB * SBA_{t-1})$	-0.0020	(0.0002) ***	0.0289	(0.0044) ***
$\ln(DEP_{t-1})$	-0.0025	(0.0015)	-0.0082	(0.0032) **
$\ln(LQ_{t-1})$	0.0254	(0.0045) ***	-0.1586	(0.0248) ***
$\ln(HPI_{t-1})$	0.0129	(0.0018) ***	-0.1699	(0.0149) ***
J-Stat	50.1631		49.6131	
P-Value (J)	0.2761		0.3313	
# of observations	976		1180	

$$Y_{i,t} = \beta_0 + \beta_1 Y_{i,t-1} + \beta_2 SBA_{t-1} + \beta_3 (SUB * SBA)_{t-1} + \beta_4 DEP_{i,t-1} + \beta_5 LQ_{i,t-1} + \beta_6 HPI_{i,t-1} + \varepsilon_{i,t} \quad (1)$$

Where $Y_{i,t}$ represents the real state-level gross product ($RSGP$) per capita or unemployment rate ($UNEMP$) of state i at time t ; SBA_{t-1} , which is the lagged SBA-7(a) guaranteed loan per job supported; SUB is a dummy variable equal to one during the 2007-2009 global financial crisis (GFC) and zero otherwise. $DEP_{i,t-1}$ is bank deposit per capita, $LQ_{i,t-1}$, is the location quotient; $HPI_{i,t-1}$ is the house price index; $\varepsilon_{i,t}$ represents a mean-zero disturbance term. All the variables in equation (1) have been converted to log values. HJ-Stat is Hansen's J statistic. P-Value (J) is the associated P-value of HJ-test. Robust standard errors clustered at the state level are shown in parentheses ** = significance at 5% level, and *** = significance at the 1% level. The data for real state GDP per capita obtained directly from BEA is only available from 1997 to 2019, so equation (1) was estimated over this sample period when the dependent variable is $\ln(RSGP_t)$.

VIII. Conclusion

The common belief about small businesses is that they are vital engines of economic growth. Consequently, promoting small business growth has become a significant concern for government officials and policymakers. Due to imperfections in credit markets, resulting from information asymmetry, small businesses often face credit rationing more frequently than larger companies. To address this issue, the Small Business Administration (SBA), which is responsible for supporting small businesses, has partnered with financial intermediaries to extend loans through its guaranteed loan programs. If SBA loan guarantees effectively reduce credit rationing for small business loans, we should observe a relationship between the measures of SBA guaranteed lending activities and overall economic performance. However, existing empirical studies have not conclusively determined whether SBA's guaranteed lending programs have a positive or negative impact on economic performance, which has motivated this current study.

Using a panel of state-level data from 1990 to 2023, we examine whether SBA lending through two distinct loan programs—504 and 7(a)—has a significant impact on state-level real GDP and the unemployment rate, and how this impact changed during the global financial crisis of 2007-2009. Employing a dynamic panel data model with a Generalized Method of Moments (GMM) estimator, we find a statistically significant positive impact of SBA lending on promoting state-level output for both loan programs. Additionally, we discover that SBA lending under the 7(a) loan program significantly reduces the state unemployment rate. Specifically, a 1% increase in SBA 7(a) loans per job supported correlates with a decrease of more than six basis points (0.06%) in the state unemployment rate. Given the steady growth of the SBA loan guarantee programs over the years, these positive impacts on economic growth and employment rates warrant continued support for the SBA loan guarantee programs. However, these positive effects diminished during the global financial crisis, resulting in a 0.002% decrease in state-level real GDP and a 0.03% increase in the state unemployment rate for both loan programs. This reduced effectiveness of SBA lending on state economic performance during the financial crisis indicates that more effort could have been made to mitigate the decline related to the crisis in state economic performance.

References

- Arellano, M., & Bond, S.R. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies*, 58(2), 277—297.
- Armstrong C., Craig B., Jackson, W.E., & Thomason, J.B. (2014). The moderating influence of financial market development on the relationship between loan guarantees for SMEs and local employment rates. *Journal of Small Business Management*, 52(1), 126–140.
- Brown JD, & Earle, J.S. (2017). Finance and growth at the firm level: evidence from SBA loans. *Journal of Finance*, 72(3), 1039–1080.
- Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1), 115-143.
- Blundell, R., & Bond, S. (1999). GMM estimation with persistent panel data: an application to production functions, IFS Working Papers W99/04, Institute for Fiscal Studies.
- Cortes, B. S. (2010). Impact of small business administration lending on state-level economic performance: A panel data analysis. *International Journal of Business and Finance*

- Research*, 4(3), 55–66.
- Craig, B. R., Jackson, W. E., & Thomson, J. B. (2007). Small firm finance, credit rationing, and the impact of SBA-guaranteed lending on local economic growth. *Journal of Small Business Management*, 45(1), 116–132.
- Craig, B. R., Jackson, W. E., & Thomson, J. B. (2008). Credit market failure intervention: do government sponsored small business credit programs enrich poorer areas? *Small Business Economics*, 30(4), 345–360.
- Craig, B. R., Jackson, W. E., & Thomson, J. B. (2009). The economic impact of the Small Business Administration's intervention in the small firm credit market: a review of the research literature. *Journal of Small Business Management*, 47(1), 221–231.
- De Ruyg, V. (2005). Are small businesses the engine of growth? American Enterprise Institute Working Paper #123.
- Driscoll, J. C. (2004). Does Bank Lending Affect Output? Evidence from the U.S. States. *Journal of Monetary Economics*, 51(3), 451–471.
- Hayakawa, K. (2009). First Difference or Forward Orthogonal Deviation- Which Transformation Should be Used in Dynamic Panel Data Models?: A Simulation Study. *Economics Bulletin*, Vol. 29(3), 2008–2017.
- Hicks, M. J. (2004). Do Rural Areas in the USA Need Microfinance? Some Early Evidence from the Community Reinvestment Act Data,” *Journal of International Banking Regulation*, 5(3), 243–250.
- Higgins, M.J. & Lacombe, D.J., Stenard, B.S., & Young, A.T. (2021). Evaluating the effects of Small Business Administration lending on growth. *Small Business Economics*, Springer, vol. 57(1), 23–45.
- Krishnan, K., Nandy, D.K., & Puri, M. (2015). Does Financing Spur Small Business Productivity? Evidence from a Natural Experiment. *Review of Financial Studies*, 28(6), 1768–1809.
- Lee, Y. S. (2018). Government guaranteed small business loans and regional growth. *Journal of Business Venturing*, 33(1), 70–83.
- Levine, R. (1997). Financial Development and Economic Growth: Views and Agenda. *Journal of Economic Literature*, 35(2), 688–726.
- Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica* 49(6), 1417–1426.
- Orzechowski, P.E. (2020). U.S. Small Business Administration loans and U.S. state-level employment. *Journal of Economics and Finance*, 44 (), 486–505.
- Roodman, D. (2009). How to do xtabond2: An introduction to difference and system GMM in Stata. *Stata Journal*, 9(1), 86–136.
- Rupasingha, A., Crown, D., & Pender, J. (2019). Rural business programs and business performance: The impact of the USDA's Business and Industry (B & I) guaranteed loan program. *Journal of Regional Analysis and Policy*, 59(August), 701–722.
- Shaffer, S. (2006). Establishment size by sector and county-level economic growth. *Small Business Economics*, 26, 145–154
- Stiglitz, J. E., & Weiss, A. W. (1981). Credit rationing in markets with imperfect information. *The American Economic Review*, 71(3), 393–410.
- Wintoki, M. B. & Linck, J.S. & Netter, J. M. (2012). Endogeneity and the dynamics of internal corporate governance. *Journal of Financial Economics*, 10, 5(3), 581–606.

