THE UNIVERSITY OF QUEENSLAND

AUSTRALIA

Asymmetric Information, Fraud, and Switching Costs in Healthcare Insurance Markets

Ahmad "Jordan" AlZu'Bi

Ph.D. Degree in Risk Management and Insurance

Masters of Investment and Finance

Bachelor of Risk Management and Insurance

A dissertation submitted for the degree of Doctor of Philosophy (Risk Management and Insurance)

at

The University of Queensland in 2021

UQ Business School

Abstract

This thesis contains three essays dedicated to the role of religion and culture in asymmetric information, fraud, and switching costs in healthcare insurance markets. In the first essay, using individual level insurance data from a Takaful health insurance company located in the United Arab Emirates (UAE), we consider how culture and religion influence asymmetric information. We find evidence of asymmetric information among religious and nonreligious individuals. However, we find that the extent of asymmetric information for religious individuals is lower than that for nonreligious individuals. Our results are similar when we consider major Muslim holidays and the Sunday effect. Finally, we provide some evidence that cultural background is also related to claim performance.

Using the same dataset in the second essay, we then identify and detect potential fraud schemes in private healthcare insurance markets. Our results reveal that fraud/abuse in the healthcare insurance market is mostly committed by healthcare providers (HCPs) and that fraudulent claims are more likely to be rejected when the HCP has control versus when the policyholder (PH) has control of the claim. Further, our results indicate that the most useful variables to identify potentially fraudulent claims are (1) the demographic characteristics of a claimant (such as age, sex, marital status, and dependency) and (2) financial characteristics (such as claim size, copays, and coinsurance).

In the third essay, we empirically examine how switching between medical plans interacts with adverse selection in the context of healthcare insurance plans. We provide important evidence that adverse selection is preset in the market and that switching costs implies an asymmetry of medical utilization between switchers, nonswitchers and those who add to their policy. More specifically, we show that the post-switching medical utilization of switchers, especially those who ADD to a current medical policy, is economically larger than the medical utilization of nonswitchers who remain in the HIGH or LOW medical plans. This suggests that the presence of adverse selection is more pronounced among switchers than among nonswitchers and is even more pronounced among those who ADD/SWITCH UP than among those who switched DOWN and stayed in the LOW/HIGH medical plans.

Declaration by the author

This thesis is composed of my original work and contains no material previously published or written by another person, except where due reference has been made in the text. I have clearly stated the contribution made by others to jointly authored works that I have included in my thesis.

I have clearly stated the contribution of others to my thesis as a whole, including the contributions they have made in statistical assistance, survey design, data analysis, significant technical procedures, professional editorial advice, and any other original research work used or reported in my thesis. The content of my thesis is the result of work I have carried out since the commencement of my research higher degree candidature and does not include a substantial part of work that has been submitted to qualify for the award of any other degree or diploma in any university or other tertiary institution. I have clearly stated which parts of my thesis, if any, have been submitted to qualify for another award.

I acknowledge that an electronic copy of my thesis must be lodged with the University Library and subject to the policy and procedures of The University of Queensland, the thesis be made available for research and study in accordance with the Copyright Act 1968 unless a period of embargo has been approved by the Dean of the Graduate School.

I acknowledge that the copyright of all material contained in my thesis resides with the copyright holder(s) of that material. Where appropriate, I have obtained copyright permission from the copyright holder to reproduce material in this thesis.

Publications included in this thesis

No publications included.

Submitted manuscripts included in this thesis

No submissions included

Contributions by others to the thesis

No contributions by others have been made to this thesis. Under the guidance of supervisors Kelvin Tan, Lori Medders, and Luke Connelly, I was responsible for all sections of the thesis.

Other publications during candidature

No other publications.

Statement of parts of the thesis submitted to qualify for the award of another degree

No works submitted towards another degree have been included in this thesis.

Research Awards and Presentations

A. In recognition of the contribution of my work, I have been awarded the following **four research awards**.

1. The first essay was awarded the prestigious "**Mark Dorfman Award**," for the "**Best PhD Paper**" at the 2020 Western Risk and Insurance Association Annual Meeting (January 2020).
2. **I was awarded the UGA Ph.D. Student Symposium Travel Award** for attending the PhD symposium at the University of Georgia (February 2020).
3. The second essay was awarded the **SWFA Doctoral Stipend Award** for presentation at the 2020 Southwestern Finance Association Annual Meeting (March 2020).
4. **I was a recipient of one of the Laney Graduate Student Council (LGSC) Travel Grants**, Emory University (August 2020).

B. These three essays have been presented at the following eleven **international and domestic conferences** since my enrollment at UQ (RQ1 2019):

1. The 2021 Asia-Pacific Risk and Insurance Association Annual Meeting, Jerusalem, Israel (accepted: July 2021).)
2. The 2021 Western Risk and Insurance Association Annual Meeting (March 2020).
3. The 2020 Virtual Southern Risk and Insurance Association Annual Meeting (December 2020).
4. The 33rd PhD Conference in Economics and Business, Monash Business School, Monash University (November 2020).
5. The 2020 Southwestern Finance Association Annual Meeting, San Antonio, Texas (March 2020)
6. The 2020 Western Risk and Insurance Association Annual Meeting, Puerto Vallarta, Mexico (January 2020).
7. The 2019 Southern Risk and Insurance Association Annual Meeting, Charleston, SC (November 2019).
8. The 32nd PhD Conference in Economics and Business, Canberra, ACT (November 2019).
9. The UQBS Finance PhD Research Day Program, Brisbane, QLD (July 2019).
10. The 2019 UQBS Research Students' Colloquium, Brisbane, QLD (July 2019).
11. The UQBS Brownbag Seminar, Brisbane, QLD (March 2019).

Acknowledgments

It has been almost six years since I started working on my PhD, which makes it the most significant milestone in my life. I have been fortunate to work closely with a large and diverse group of faculty members, visiting scholars, and PhD students. I'm forever grateful for all the help and support I have received during my tenure as a PhD student, especially from my supervisors and my family members. Without all of you, earning my doctoral degree would have been impossible.

A big "Thank You" to my dissertation advisory team: Associate Professor Kelvin Tan (principal advisor), Professor Lori Medders (associate advisor), and Professor Luke Connelly (associate advisor). Under Kelvin's priceless supervision and academic support over the last three years, I have not only gained intellectual knowledge but also expanded my theoretical and empirical base to a wide variety of academic literature in finance and insurance. Kelvin, you will always be my role model and a big inspiration, and I will always support my students the way you have supported me. Lori and Luke, your dedicated supervision has also enhanced my research/teaching experience. Your valuable feedback and suggestions are sincerely appreciated.

I would like to thank my reader, Associate Professor Barry Oliver, for the time he has devoted to the examination process of my dissertation. My research has been enhanced as a result of his feedback and suggestions. Finally, special thanks to my host supervisor at Emory University, Professor David Howard, for his support and guidance during my visit at Emory. David has been very supportive and accommodating, especially during the COVID-19 pandemic.

I'm forever thankful to my family members for their unconditional love and support (Um Samer, Abu Samer, Samer, Mohammad, Ali, Hiba, Samir, Mahmoud, Omar, Abdallah, Yaurob, Amer, and Afnan). All of you have always believed in me and supported me emotionally and financially during tough times, especially Samer, Mohammad, and Ali. My work is dedicated to all of you, and I hope I have made you proud.

I was lucky to meet many PhD students during my PhD candidature, especially during my time at the UQ Business School and Emory University. I have become a close friend to many of you, especially Trinh Hue Le, Sajeeb Saha, Mahsa Amirzadeh, Nicole Hay, Yasmine Azzaoui, Michael Flores, Viobillyn "Lyn" Lapore, Vanvilay "Kop" Phommalath, and Shresht Shetty. Your friendship has been a significant part of my PhD journey. To my Vietnamese friend, Trinh Hue Le: the cheerful interaction with you has always made me more focused on finishing what I have started. You filled my PhD life with joy and happiness. I will miss sharing a workplace with you.

Last, I am grateful for the Australian Government Research Training Program Scholarship, the UQ Business School Research Scholarship, and the Emory University Tuition Scholarship programs,

which have supported me financially, providing me with numerous research/conference scholarships during my PhD candidature. This full financial support allowed me to focus on PhD coursework, research, and teaching rather than to be constantly consumed with worry over finances.

Certainly, my PhD journey has been a learning curve and will never stop. In Arabic, we say, “The journey of a thousand miles begins with one step,” and earning my PhD degree is just the starting point of another journey. I look forward to starting my favorite career as a professor of finance, risk management, and insurance.

Ahmad Aqil Ayed AlZu’Bi
Monday 11/1/2021

Financial Support

This research was supported by the Australian Government Research Training Program Scholarship, UQ Business School Research Scholarship, and Emory University Tuition Scholarship.

Keywords

Asymmetric Information, Adverse Selection, Moral Hazard, Insurance Economics, Insurance Fraud, Healthcare Insurance, Takaful Insurance, Culture, Religion, and Health Economics

Australian and New Zealand Standard Research Classifications (ANZSRC)

ANZSRC code: 150204, Insurance Studies, 60%

ANZSRC code: 150201, Finance 20%

ANZSRC code: 140208, Health Economics 20%

Fields of Research (FoR) Classification

FoR code, 1502, Banking, Finance and Investment, 80%

FoR code: 1402, Applied Econometrics, 20%

Table of Contents

List of Tables

List of Figures

Chapter 1: General Introduction

This thesis contains three essays dedicated to the role of religion and culture in asymmetric information, fraud, and switching costs in healthcare insurance markets.

The **first essay** investigates the impact of religion and culture on asymmetric information in the Takaful (Islamic) healthcare insurance market. The current insurance literature has focused mainly on documenting whether asymmetric information is present across all insurance lines but has not focused on the factors and circumstances that are likely to influence the presence of asymmetric information and selection (Cohen and Siegelman 2010). This essay contributes to the literature by answering the call of Cohen and Siegelman (2010) to direct the focus of the insurance economists' future research towards the drivers and contributors of asymmetric information. To the best of our knowledge, this is the first study that considers how religion and culture might influence asymmetric information by examining the presence of asymmetric information within the specialized Takaful health insurance market.

We find that the correlation between risk and claim performance (a proxy for information asymmetry) for Muslim policyholders (PHs) is lower than that for non-Muslim individuals. However, the risk-claim-performance results are sensitive to the definition of claim performance. My empirical results are consistent with my theoretical model, in which we show that religiosity is a key factor in the demand for Takaful insurance and can be a significant driver of a reduction in negative behaviors that influence the presence of asymmetric information. Furthermore, when we control for major Muslim religious holidays (e.g., Eid Al-Fitr, Ramadan, and Eid Al-Adha) and the Sunday effect in insurance markets, the results are robust and consistent with the theoretical model. Finally, we provide evidence that cultural background is also related to claim performance.

The **second essay** aims to identify fraud patterns in healthcare insurance markets and then examines whether and to what extent fraud is committed largely by healthcare providers (HCPs) rather than by healthcare insurance PHs. The results reveal that fraud/abuse in the healthcare insurance market is committed mostly by HCPs. Furthermore, fraudulent claims are more likely to be rejected when the HCP rather than the PH controls the claim. Our results also indicate that the most useful variables in identifying potentially fraudulent claims are (1) the demographic characteristics of a claimant (such as age, gender, marital status, and dependents) and (2) financial characteristics (such as claim size, copayment, and coinsurance).

The **second essay** makes the following two contributions. First, to the best of our knowledge, it is the first study to identify and detect fraud schemes in two private health insurance markets, namely, the

Takaful insurance market and the commercial insurance market. Specifically, it is the first paper to provide empirical support that healthcare fraud is more pronounced among HCPs than among PHs. Second, this paper contributes to the existing literature on insurance fraud by using a large sample of claims to identify and detect potential fraud schemes (the dataset contains 633,042 claims versus 1,350 claims in Artis, Ayuso, and Guillen (2002)).

The **third essay** examines the relationship between switching costs and adverse selection in a setting of endogenous pricing in the context of Takaful (Islamic) healthcare insurance, where religion and culture are crucial drivers of demand for Takaful products. Specifically, we empirically examine how switching between medical plans interacts with adverse selection in the context of healthcare insurance plans. In contrast to prior research, we further consider how adding to current medical policy (instead of switching UP or DOWN) affects the extent of adverse selection. The work controls for the presence of moral hazard by construction because we compare medical utilization between switchers and nonswitchers within the same medical plans.

The **third essay** reveals important evidence that adverse selection is present in the market and that switching costs implies asymmetry of medical utilization among switchers, nonswitchers and those who add to their policy. Specifically, we show that the post-medical utilization of switchers, especially those who ADD to their current medical policy, is economically larger than that of nonswitchers who remain in HIGH (i.e., generous) or LOW (i.e., economy) medical plans.

As a larger dataset from the Takaful healthcare insurance market was used, this essay contributes to the insurance and health economics literature that examines adverse selection in healthcare insurance markets (e.g., Polyakova, 2016; and Cohen and Siegelman 2010) and the determinants of a PH's switching behavior (e.g., Cardon, 2018; Handel, 2013; Cutler et al., 2010; Tchernis et al., 2006; Cutler and Zeckhauser, 2000; and Altman et al., 1998). The dynamic relationship between switching costs and adverse selection has never been explored in the Takaful insurance markets. Furthermore, the current literature has yet to examine the dynamic relationship between policyholders (PHs)' adding behavior and their medical utilization. To the best of our knowledge, this is the first study to examine switching costs and evaluate the potential interaction between switching behavior, adding, and adverse selection in the private Takaful health insurance market, where employees and their dependents are covered by employer-sponsored health insurance.

2. Chapter 2: Asymmetric Information and Culture: Evidence from Takaful Health Insurance

2.1 Introduction

The existence of asymmetric information and adverse selection in many types of conventional insurance has been widely documented (Cohen and Siegelman, 2010). However, there has been no focus on whether asymmetric information and/or adverse selection are present in the Takaful (Islamic) insurance market and whether religion and culture influence the extent of asymmetric information and/or adverse selection across different lines of insurance. For example, Cohen and Siegelman (2010) argue that future research on asymmetric information should not focus on documenting whether adverse selection is present but rather focus on the factors and circumstances that are likely to influence the presence of adverse selection. In this paper, we answer this call by examining how religion and culture's association with behavior affects asymmetric information in the Takaful insurance market. Furthermore, we develop a theoretical model to capture the role of religiosity in the Takaful insurance demand and its effect on negative drivers of selection and asymmetric information (e.g., fraud, misrepresentation, and withheld information) in the healthcare insurance market.

Religion and culture are potential significant drivers of policyholder behavior because the insurance policyholders' insurance demand is based not only on their rational decisions but also on their cultural perceptions and religious beliefs (Outreville 2018, Trinh, Nguyen, and Sgro 2016, and Park and Lemaire 2012). Furthermore, prior behavioral insurance literature focusing on culture and insurance demand reveals that Islamic laws (e.g., Sharia law) negatively influence insurance demand, especially within Muslim countries (Outreville 2018, Trinh, Nguyen, and Sgro 2016, and Park and Lemaire 2012).

The results of this work are relevant and important because Takaful insurance is growing in the US, Europe, and other parts of the world (Kwon, 2007)[1]. Most importantly, the Global Takaful Market Size, Status and Forecast 2019-2025 report shows that global gross Takaful contributions reached USD 22.5 billion in 2018 and are expected to grow to 49.5 billion by the end of 2025, with a compound annual growth rate of 11.65 percent (see Appendix A). Furthermore, there are 80 Takaful insurance companies around the world, with an additional 200 "Takaful windows." Moreover, Kwon (2007) shows that "Muslims are ubiquitous, and the population continues to rise," currently representing approximately 25 percent of the world's population. The most updated statistics regarding Muslim populations in 2014

[1] For instance, in Europe, the Takaful market size was USD 61.7 million in 2019 and is expected to reach USD 140 million by the end of 2025. However, in the US alone, the Takaful market size was USD 10.3 million in 2019 and is expected to reach USD 21.3 million by the end of 2025. See Figures A2, and A3 in Appendix A: Global Takaful Market Size/Share by Regions.

reveals that Islam is the fastest-growing religion worldwide and has obtained the largest global following of all religions.[2] Given these statistics, it is worth exploring the association between religion, culture, and asymmetric information in the Takaful insurance market.

Sharia law is likely to influence asymmetric information in insurance markets because behaviors such as lying, misrepresentation, fraud, and withholding information are major contributors to asymmetric information and are prohibited by Sharia (Rahman and Daud 2010). Furthermore, Sharia law forbids Muslims from buying commercial insurance or any non-Islamic insurance. As Kwon (2007) notes, the law forbids engaging in activities that involve ambiguity or uncertainty as well as vague contracts in which the outcomes are hidden or speculative in nature. In addition, according to such Islamic laws, misrepresentation, fraud, withholding information, and lying are rejected in all forms[3], and these tenets of Islamic laws are similar to those of other religions, particularly those religions that also adhere to the concept of a judging deity. This belief suggests that religious policyholders may be less likely to misrepresent themselves, commit fraud, withhold information and/or misuse insurance because of their religious beliefs. Furthermore, this belief may suggest that religion and culture are influential factors when insurance underwriters require potential policyholders to provide information about themselves and/or when current policyholders (mis)use their insurance in claims filings.

Our results are consistent with our expectation, and we provide evidence of significant asymmetric information in the Takaful health insurance market. The results reveal that the correlation between risk and claim performance (a proxy for information asymmetry) for Muslim policyholders is lower than that for non-Muslim policyholders. More specifically, the results suggest that Muslims file fewer claims than non-Muslims and that Muslims' average claim size is also smaller than that of non-Muslims. We consider this finding as evidence of asymmetric information for both religious and nonreligious individuals. Our results are also evident and consistent with our theoretical model when we control for major Muslim religious holidays (e.g., Eid Al-Fitr, Ramadan, and Eid Al-Adha) and the Sunday effect in insurance markets. The results show that asymmetric information and excessive claiming behavior are significantly lower during the Muslim holidays and Sundays (after the weekend), suggesting that the Sunday effect is

[2] The population information is based on worldwide data on the Muslim population [www.islamicpopulation.com]. According to the 2014 statistics, the Muslim population stood at 2.08 billion, representing approximately 28.26 percent of the world's population—an increase from approximately 25 percent since 2007.

[3] Prophet Muhammad (peace be upon him) emphasized in various occasions that Islam rejects lying and deception in all forms, and said, "You must be truthful, for truthfulness leads to righteousness and righteousness leads to Paradise. A man will keep speaking the truth and striving to speak the truth until he will be recorded with Allah as a siddeeq (speaker of the truth). Beware of telling lies, for lying leads to immorality and immorality leads to Hellfire. A man will keep telling lies and striving to tell lies until he is recorded with Allah as a liar."

present and that the reduction in asymmetric information is pronounced even during Muslim religious holidays and Sundays. Finally, we find some evidence that cultural background is also related to claim performance. However, some of the results are economically insignificant. Given this evidence, we feel that these findings require further attention by insurance economists.

This paper allows us to contribute to the broader literature on the relationship between religion, culture, and economic behavior. More specifically, this paper contributes to the existing literature by examining the presence of asymmetric information within a specialized Takaful health insurance market for a large sample of policyholders with different religions, nationalities and cultural backgrounds. Furthermore, given the uniqueness of our data, which contain the religion and nationality for every single policyholder in our sample, our paper can answer Cohen and Siegelman's (2010) call for literature to focus on the factors and circumstances that are more likely to influence the presence of asymmetric information. To the best of our knowledge, our paper is the first to consider how religion and culture might influence asymmetric information and the first to theoretically explain the role of religiosity in Takaful (Islamic) insurance demand and its potential effect on negative behaviors that influence the presence of asymmetric information[4]; prior empirical work, such as Outreville (2018), Trinh, Nguyen, and Sgro (2016), Park and Lemaire (2012), and Chui and Kwok (2008), focuses only on how national culture influences insurance demand across various countries, including some Muslim countries. We can expand on their analysis by using a unique dataset of individuals who reside in the same country and therefore live under a similar institutional setting but have different cultural backgrounds. More specifically, we extend the existing literature by considering the relationships between religion, culture, and insurance utilization (measured by claim behavior) by comparing the different behaviors of individuals from different cultural and religious backgrounds while holding the country factors (such as macroeconomic environments) constant. Our contribution also resonates with the call from Crawford, Pavanini, and Schivardi (2018), who state that "… deepening our understanding of the extent and causes of asymmetric information is key for the design of a regulatory framework that limits their negative consequences".

The remainder of this paper will proceed as follows. The next section provides a theoretical model for how religion may influence asymmetric information and the demand for Takaful insurance. The third section offers a detailed theoretical background and development of our hypothesis. The fourth section

[4] See Appendix C: The role of religiosity in Takaful "Islamic" insurance demand and asymmetric information.

provides a description of the data and methodology used to test our hypotheses. The fifth section presents our results, and the final section offers concluding remarks and suggestions for future research.

2.2 What is *Takaful* Insurance? Is Takaful Insurance More Expensive than Commercial Insurance?

Takaful is an Islamic law-compliant insurance in which participants (policyholders) contribute money (premiums) as donations via systematic pooling to support each other against loss or damage. Takaful policies cover health, life, and general insurance needs. Hussain and Pasha (2011) state that Takaful insurance products are Islamic alternatives to commercial insurance products. A group of policyholders agrees to jointly indemnify a loss or damage, promoting social solidarity and cooperation among members. The ultimate objective of any Takaful insurance contract is to have sufficient funds as a buffer against any potential losses for the members who are involved based on solidarity. According to QY Research[5], Takaful is different from commercial insurance in the four following major ways.

(1) Takaful is branded insurance based on Sharia or Islamic religious law, and the funds (premiums) are contributed by donations from participants to protect other participants from risk. Maysami and Kwon (1999) report, "…The Islamic model of insurance policy is based on the fundamental principle of mutual cooperation and solidarity, as ordained by Allah (SWT) mentioned to this effect in the Holy Quran."

(2) Takaful operators are required to comply with Islamic laws and invest the collected funds (premiums) in investments that are free from usury "Riba", gambling, and uncertainty[6]. Wahab et al. (2011) and AL-Amri (2013) state that people who are not familiar with Takaful insurance may have the misconception that profit-oriented business transactions are not allowed based on Sharia law. In fact, investors are allowed to engage in business transactions as long as they do not specify a predetermined return, denoting a situation occurring mainly when the outcomes of such businesses are uncertain, which could justify why conventional commercial insurance is not permissible in Islam.

(3) Unlike that of businesses invested in by commercial insurance companies, in accordance with Sharia law, the profit (surplus) of a Takaful business must be shared among policyholders and Takaful

[5] Global Takaful Market Size, Status and Forecast 2019-2025, QY Research. We purchase this report from QY Research. For more information on QY's research, please refer to its website at www.qyresearch.com.
[6] Commercial insurance companies invest in financial instruments such as bonds, stocks, options, future/forward contracts, etc. According to Islamic law, such investments are not free from usury "Riba", gambling, or uncertainty.

insurance companies (operators of a Takaful fund). Maysami and Kwon (1999) mention that profit maximization and generation are not the core goals of Takaful insurance. Furthermore, Takaful companies must share any generated profits (surplus) with participating policyholders. In addition, as halal service providers, Takaful insurance companies are required to be conservative and invest only in projects that do not violate Sharia law. This explains whay Takaful is more expesnive than commercial insurance.

(4) Takaful insurance is significantly more expensive than commercial insurance. According to the Global Takaful Market Size, Status and Forecast 2019-2025 report, Takaful insurance was approximately 36 percent more expensive than commercial insurance across all insurance lines in 2019. Table A1 shows the price differential between Takaful and commercial insurance products for 2017 to 2019. For instance, Table 3 reveals that in 2019, Takaful insurance products were USD 112 (36 percent) more expensive than comparable commercial insurance products.[7] The results hold within the healthcare insurance market (both group and individual healthcare markets), for which Table A2 shows the price differential between Takaful healthcare insurance and commercial healthcare insurance over the past 3 years (2017 to 2019). Table A2 reveals that in 2019, group Takaful health insurance products were USD 161 (35 percent) more expensive than comparable group healthcare commercial insurance products. Furthermore, individual Takaful health insurance products were USD 306 (33 percent) more expensive than comparable individual commercial health insurance products. Interestingly, the table reveals that individual healthcare insurance coverage for both Takaful and commercial insurance was significantly more expensive than group healthcare coverage.[8] The results are consistent with those of Diamond (1992), who reports that individual and small group health insurance plans are more expensive than group health insurance plans and require greater cost-sharing.

(5) The way Takaful insurance works (e.g., underwriting) is similar to how commercial insurance works. However, Takaful insurance is based on Sharia law (Islamic laws) and explains how participating policyholders are responsible for cooperating with and protecting each other. Thus, it's a product that is offered for sale to all customers regardless of their religious affiliation (see Kwon 2007 and Maysami and Kwon 1999). This may explain the high premium for Takaful. For instance, Coolen-

[7] This finding is consistent with the findings of Coolen-Maturi (2013), who report that Takaful insurance products are more expensive than their commercial counterparts and that moderate Muslims in the UK are willing to demand Takaful insurance products only if their prices and coverages are comparable to those of conventional insurance products.

[8] For instance, in 2019, Takaful insurance coverage was USD 617 (99 percent) more expensive for an individual than for the individual's Takaful group counterparts.

Maturi (2013) state that Takaful insurance is a product, and is more expensive than the commercial insurance, and some Muslims are willing to purchase Takaful products if only its services, and prices are competitive to the commercial ones.

Takaful insurance has been available since the late 1970s and was first introduced in Sudan, when a group of Islamic insurance companies were established to offer a unique insurance product for individuals and businesses (World Takaful Report, 2016). Since then, the Takaful insurance market has grown significantly and is expected to experience exceptional growth by 2025. According to the Global Takaful Market Size, Status and Forecast 2019-2025 report, the global Takaful market size was USD 22.5 billion in 2018 and is expected to grow to USD 49.5 billion by the end of 2025, with a compound annual growth rate of 11.65 percent. This growth is shown in Figure A1, which presents the global Takaful market size and its expected growth rate for the years 2014 to 2025.

Takaful arrangements are largely offered in various countries with major Muslim populations (i.e., Gulf Cooperation Council (GCC) and Southeast Asia). Remarkably, Takaful arrangements are also found in countries without large Muslim populations, such as China and countries in North America, Europe, Central America, and Sd South America. Table A3 is reproduced from the world Global Takaful Market Size, Status and Forecast 2019-2025 and presents the global Takaful market size by region for the years 2019 to 2025. Table A3 shows that the GCC and Southeast Asian countries (i.e., Saudi Arabia, Indonesia and Malaysia) dominate the global Takaful industry, exhibiting approximately 92 percent of the Takaful market share in 2019 (see Figures A2 and A3 in Appendix A: Global Takaful Market Size/Share by Regions), and their Takaful market is expected to exhibit strong growth by 2025 (12.39 percent in the GCC countries and 5.67 percent in Southeast Asia). Remarkably, Table A3 reveals the overall global contributions of Takaful in Europe, the US, China, and Central and South America. For instance, in Europe, the global Takaful market size was USD 61.7 million in 2019 and is expected to reach USD 140 million by the end of 2025, with a compound annual growth rate (CAGR) of 14.63 percent during 2019-2025. However, in the US, the global Takaful market size was USD 10.3 million in 2019 and is expected to reach USD 21.3 million by the end of 2025, with a CAGR of 12.87 percent during 2019-2025.

Overall, according to the proposals by insurance scholars[9], there are three main problems with conventional insurance (according to Sharia law). First, commercial insurance violates the prohibition of *"Gharar"* (uncertainty). Since there is a time lag between the time of buying insurance and claim occurrence, future benefits rely on the outcome of future events that are not known at the time of the

[9] See Kwon (2007), and Maysami and Kwon (1999).

signing of the contract. Furthermore, the insurance company collects premiums in exchange for indemnity against risks that may not occur. Second, insurance is based on *"Maysir"* (gambling/speculation); policyholders are essentially betting their money by paying premiums to insurers. If the policyholder experiences a loss, the insurance company will then make a payment as a result of the occurrence of that event. If the loss never occurs, the insurance company keeps the premium.[10] Third, all insurance companies have investments and saving strategies. Insurance companies essentially invest part of the premiums on behalf of their policyholders. Insurance companies conduct their business by investing the collected premiums in different projects that violate Sharia law. For instance, conducting business that includes interest-based investment *"Riba"* (usury) and/or involves any alcoholic products is forbidden. Table A4 in the Appendix summarizes the core differences between Takaful and commercial insurance[11].

2.3 Theoretical Framework: The Role of Religiosity in Takaful "Islamic" Insurance Demand and Asymmetric Information.

Insurance economists have theoretically explained insurance demand across different insurance lines.[12] However, Takaful "Islamic" insurance demand has never been explained theoretically. In this section, we construct a simple theoretical model to explain the role of religiosity in Takaful "Islamic" insurance demand and its potential effect on negative behaviors (especially fraud, misrepresentation, and withheld information) that influence the presence of asymmetric information.

The model's construction begins with the following nine assumptions:

1. The individuals' preferences for state-contingent net wealth is represented by risk aversion, which is denoted by a von Neumann-Morgenstern utility function U. U is twice continuously differentiable, with $U' > 0$, and $U'' < 0$. Individuals maximize the expected utility EU.
2. A risk-averse individual, having initial wealth, W, may experience a loss, L, where $L \geq 0$, subject to a probability of loss, p, where $0 < p < 1$ and pL represents the expected loss.
3. To isolate the role of religion on insurance demand, we assume homogenous risk factors across policyholders and assume that insurance buyers can be categorized as religious, *R*, or nonreligious, *I*, and face an equal probability of loss (p). Religious buyers may adhere to any

[10] AL-Amri (2013) states, "…The amount insured is paid to the policyholder when certain events occur. If the event never occurs, the insurance company keeps the premium."

[11] Table 5 is reproduced from the world Global Takaful Market Size, Status and Forecast 2019-2025 Report by QY Research.

[12] To view examples, see the theoretical work in healthcare insurance (Phelps, 1973), in life insurance (Yaari, 1964 and 1965), and various general theoretical models of insurance demand (Schlesinger, 2013).

religious faith (Muslim, Jewish, Christian, etc.). Since the primary emphasis of the insurance purchase, however, is the decision regarding the purchase of Takaful insurance specifically, our expectations of an individual policyholder centers on whether the policyholder is a religious (adhering) or nonreligious (nonadhering) Muslim. Thus, religiosity with respect to the Muslim faith—whether religious, $R_{M,}$ or nonreligious, I_M—is the distinction utilized in the remainder of the insurance purchase discussion.

4. We assume that the doctrinal adherence of a religious Muslim buyer is a perfect representation of another religious non-Muslim buyer. For instance, we assume that the doctrinal adherence of religious Christians or Jews is similar to the adherence of religious Muslims. The rationale behind this argument is that a (adhering) Christian, for instance, breaks no Christian law by purchasing Takaful insurance. Indeed, a religious Christian might find Takaful insurance to be superior to commercial options available because of the principles by which it operates.[13]
5. We assume that all buyers, whether they are religious or nonreligious, have full access to Takaful and commercial insurance products. For simplicity, we assume that insurance is available at an affordable and adequate price for both Takaful (V^T) and commercial insurance (V^C)products and that for both products, the price of insurance is increasing in L (V^T and V^C, respectively) such that W> V^T and W>V^C (affordable) and $V^T \geq pL$ and $V^C \geq pL$ (adequate).
6. Buyers can choose no insurance or some amount of insurance, α, where $0 \leq \alpha \leq 1$, with respect to the amount of loss, L. There is no insurance mandate that constrains the buyer's decision to a purchase of a minimum level of α.
7. We further assume that an individual's decisions hold a nonmonetary (intrinsic) value, **Ω**, most easily described here as the subjective value ascribed by an individual to religious adherence, personal integrity and reputation value. Islamic laws forbid the purchase of commercial insurance (Kwon 2007); therefore, for example, religious Muslims receive +**Ω** by following Islamic laws (by demanding only Takaful insurance).[14] For a religious Muslim, there is not only an intrinsic gain from purchasing Takaful insurance over purchasing commercial insurance but

[13] The acceptance, tolerance, and similarities between religions can explain the similar behavior of people of faith. For instance, Ayoub (1991) explains that the acceptance and tolerance between Christianity and Islam and states that "… while in both religions, tolerance is a fundamental principle based on the imperative of love and respect for human life and dignity, the Quran clearly advocates mutual acceptance and cooperation among the people of the Book: Jews, Christians and Muslims. This is evidenced in the term ahl al-kitab, the family of the Book, which includes all the children of Abraham.

[14] The religious value represents the deeds (credit) a Muslim would earn from God "Allah" by not violating the Islamic laws. According to Sharia laws and Prophet Mohammad's teachings "Sunnah", such deeds move good Muslims closer to the heaven. Most importantly, such deeds have no economic value to adhering Muslim buyers but comprise a tremendous religious value. For a religious Muslim, this value is significantly larger than the amount of money he/she would save due to price differential between the Takaful option and the comparable commercial insurance products.

also an intrinsic gain by purchasing Takaful insurance in lieu of going uninsured, since under Islamic law, Allah rewards "good" decisions. Alternatively, this nonmonetary value is negative ($-\Omega$); i.e., it denotes a cost, if a religious Muslim violates Islamic laws (i.e., purchases commercial insurance).[15] Table 2.1 summarizes the valuation of Ω within the insurance purchase decision for an adhering Muslim.

Table 2.1 Intrinsic Value Effect of Insurance for Adhering Muslim Buyers

No insurance	Takaful insurance	Commercial insurance
$\Omega = 0$	$\Omega \geq 0$	$\Omega \leq 0$

8. Takaful insurance holds a nonmonetary value, Ω, only for religious Muslims since all others do not ascribe intrinsic importance, or value, to Sharia law. Thus, $\Omega = 0$ within the insurance purchase decision for a nonreligious Muslim or for a non-Muslim.
9. Coolen-Maturi (2013) reports that Takaful insurance products are more expensive than their commercial counterparts.[16] Given that the purchase of commercial insurance violates Islamic law and that the insurance decision is observable by others, it is intuitively appealing that for religious Muslims, the utility value of Ω (whether owing to its adherence, integrity or reputation elements), $U(\Omega)$, is sufficiently large such that it creates a relatively inelastic demand for Takaful insurance compared to a commercial insurance substitute; thus, the utility value of the adverse price differential, $U(V^T > V^C)$, does not materially enter the religious Muslim's insurance decision.[17] Thus, for a religious Muslim, $U(V^T > V^C) = 0$.

Assuming that all policyholders are rational and demand the option that optimizes their overall utility function, we expect religious Muslim buyers to buy no insurance or to gain sufficient utility from the intrinsic value of the product (*Halal*) such that they will demand some amount of Takaful insurance, regardless of the availability of otherwise attractive commercial insurance products (*Non-Halal*). Nonreligious Muslims and non-Muslims each have utility functions based on economic value and demand the least expensive comparable products. This expectation is supported by a Coolen-Maturi

[15] According to Islamic laws, "…bad deeds cancel out some good deeds, but they may be restored if one repents" Fath al-Baari by Ibn Rajab (1/146).
[16] The Global Takaful Market Size, Status and Forecast 2019-2025 report revealed that 2019 Takaful insurance products were USD 112 (36 percent) more expensive than comparable commercial insurance products. The results held within the healthcare insurance market (both group and individual healthcare markets). For more information, see Appendix A.

(2013) survey, which indicates that moderate Muslims in the UK are willing to demand Takaful insurance products only if their prices and coverages are comparable to conventional insurance products.
An individual's baseline, or reservation, expected utility, denoted by EU^0, can be characterized simply as follows:

$$EU^0 = p\,U(W\text{-}L) + (1\text{-}p)\,U(W), \tag{1}$$

where $0 < p < L$ and $L > 0$. U^0 implies no insurance; thus, with p chance of occurring, an individual is subject to a loss L from initial W. Regardless of the individual's religion or level of religiosity (religious adherence), the expected utility of having no insurance is driven by the satisfaction derived from their expected wealth, given that the individual may or may not suffer a loss to their wealth. The no-insurance case holds no intrinsic value, either positive or negative, $\pm\,\boldsymbol{\Omega}$, for the religious Muslim, R_M, or for the irreligious Muslim or non-Muslim, I_M.

2.3.1 The Case of the I_M Buyer (the Irreligious Muslim or Non-Muslim)

For the risk-averse yet irreligious Muslim, $\Omega = 0$ within the insurance decision. Only if $V^T = V^C$ will the irreligious Muslim purchase Takaful insurance over commercial insurance since a disutility $(U(V^T\text{-}V^C))$ results otherwise from the purchase of Takaful insurance. Therefore, for $V^T > V^C$, the insurance choice for this buyer is effectively a decision between commercial insurance and no insurance.

For this individual, when $V^T\text{-}V^C \geq 0$, the expected utility is maximized as follows:

$$\text{Max } EU_I = (1\text{-}p)\,U(W - V^C) + p\,U(W\text{-}\,V^C - L + \alpha L), \tag{2}$$

where $W > 0$ and $V^C > L > 0$.

Solving for the first-order conditions with respect to α yields

$$\alpha^*:\; -(1 - p)\,pL \bullet U'(W - V^C) + p[U'(W - V^C - L + \alpha L) \bullet (L - pL)] = 0 \tag{3}$$

Rearranging Equation 3, the optimal amount of insurance, α, is chosen where:

$$p\,[U'(W - V^C - L + \alpha L) \bullet (L - pL)] = (1 - p) \bullet pL \bullet U'(W - V^C). \tag{4}$$

Reducing this result, and as long as $V^C = \alpha pL$ (actuarially fair price), full insurance coverage will be the optimal insurance choice. Indeed, even with $V^C > \alpha pL$, the expected utility from buying full insurance ($\alpha = 1$) is greater than the expected utility of having no insurance ($\alpha = 0$) so long as the marginal utility from the insurance benefit, $pU'(\alpha L)$, is greater than the marginal disutility of the price paid for the insurance, $U'(V^C)$.

2.3.2 The Case of the R_M Buyer (the Religious Muslim)

Now suppose that by purchasing some amount of Takaful insurance at price V^T, a religious Muslim can insure against the loss (L) and gain the utility of the religious, nonmonetary value (+Ω) by adhering to Sharia law. The risk averse, religious Muslim's decision regarding Takaful insurance can be characterized by

$$\text{Max } EU_R = (1\text{-}p)\, U(W - V^T + \Omega) + p\, U(W\text{-}\ V^T - L + \alpha L),^{18} \qquad (5)$$

where Ω>0 because the individual has chosen against violating Islamic law (i.e., did not purchase commercial insurance), W > 0, and $V^T > L > 0$.

Solving for the first-order conditions with respect to α yields

$$\alpha^*:\ -(1 - p)\, pL \bullet U'(W - V^T + \Omega) + p\, [U'(W - V^T - L + \alpha L) \bullet (L - pL)] = 0 \qquad (6)$$

Rearranging Equation 3, the optimal amount of insurance, α is chosen where:

$$p[U'(W - V^T - L + \alpha L) \bullet (L - pL)] = (1 - p) \bullet pL \bullet U'(W - VT + \Omega). \qquad (7)$$

As long as $V^T = \alpha pL$ (actuarially fair price), full insurance coverage will be the optimal insurance choice. Indeed, even with $V^T > pL$, the expected utility from buying full insurance (α = 1) is greater than the expected utility of having no insurance (α = 0) so long as the marginal utility from the insurance benefit and intrinsic, religious value, p[U'(αL) + Ω], is greater than the marginal disutility of the price paid for the insurance, $U'(V^T)$.

Table 2.2 summarizes the decision conditions under which religious versus nonreligious Muslims would buy Takaful or commercial insurance. Table 2.2 shows that a religious Muslim always demands Takaful insurance across all market conditions. A nonreligious Muslim, however, demands the least expensive comparable insurance coverage and is indifferent between the two if their prices and coverages are comparable[19].

---

[18] Because of the price differential, $V^T – V^c$, given that Takaful products are more expensive than the comparable commercial insurance products, one might expect the religious Muslim's utility function to include the disutility, $-V^T > V^C$. As stated in the basic model assumptions, however, religious Muslims care significantly more about the religious, nonmonetary value +Ω than the monetary value $v(L)^T – v(L)^C$; therefore, $v(L)^T – v(L)^C$ is not materially important and is not represented in Equation 2. From a religious Muslim's point of view, the intrinsic utility of the religious value (Ω) is infinitely larger than the financial utility he/she would gain by buying a cheaper commercial insurance product. From a traditional economic perspective, this may be a surprising (unexpected) assumption and notable for that reason.

[19] Note that in the healthcare insurance markets, irreligious non-Muslims may also choose to buy Takaful insurance due to the following two reasons: (1) if the majority of the group members are inclined to buy Takaful insurance, an irreligious non-Muslim might feel the peer pressure to join in the Takaful insurance purchase; (2) irreligious non-Muslims may make a

Table 2.2. Insurance Choice and Decision Conditions for Religious versus Nonreligious Buyers

Insurance Choice	Market Condition	Religious Muslim Buyer R_M	Nonreligious Muslim Buyer I_M
Takaful insurance, commercial insurance, or no insurance	$v^T > v^C \geq \alpha pL$	Always demands Takaful Insurance (No religious incentives to buy commercial insurance)	Never demands Takaful (No economic incentives to buy Takaful insurance)

The key question is, if following Islamic laws and earning a large, positive nonmonetary value is a major driver of the demand for Takaful insurance by religious Muslims, is it also a significant driver of policyholder underwriting and claims behavior? If so, to what extent does this influence the level of asymmetric information in the Takaful insurance market? These questions are discussed in the next section.

2.3.3 Do Islamic Laws Influence Policyholder Behaviors?

Section 2.3.2 introduces the importance of the nonmonetary value ($\pm\, \Omega$) in purchasing Takaful insurance products for religious Muslims. We assume that this value has no extrinsic economic value, as it represents simply the intrinsic value of religious adherence, personal integrity and reputation (including for religious Muslims, the good deeds credit a Muslim would earn from *Allah* due to following Islamic laws). Such nonmonetary values impact the religious individuals' utility functions and can result in a disutility if one engages in a decision or behavior that compromises one's religious doctrine, personal integrity or reputation.

While buying Takaful insurance may optimize a religious Muslim's utility function, dishonest behaviors—misrepresentation, fraud, withheld information, and lying—are rejected in all forms according to Islamic laws; thus, these negative behaviors enter a religious Muslim policyholder's utility function as a disutility. The question is, if religious Muslims care about earning positive nonmonetary

rational decision, as buying healthcare insurance with the large group is cheaper than purchasing different coverage by creating a small group or buying it from the individual insurance markets. Diamond (1992) reports that individual and small group health insurance plans are more expensive than group health insurance plans and require greater cost sharing. The same argument applies to non-Muslims who buy Takaful insurance products.

value (positive deeds), to what extent do such intrinsic values influence the negative drivers of asymmetric information, such as those listed above, in the Takaful insurance market?

To answer these questions, we start similarly as in our earlier Takaful insurance demand model and assume the following.

1. A risk averse buyer decides to purchase insurance. Based on the insurance purchase model developed in Section 2.1 and assuming otherwise comparable insurance product choices, a religious Muslim always buys Takaful insurance, and an irreligious Muslim or non-Muslim buys commercial insurance. An irreligious Muslim or non-Muslim who purchases Takaful insurance perceives there is no comparable commercial insurance product available.
2. Risk averse policyholders are presented with an opportunity to behave dishonestly either through risk misrepresentation (underwriting) or insurance fraud (claims). Insurance companies utilize pre-loss underwriting and claims auditing to catch misrepresentation and fraud. Dishonest behavior is subject to a low probability (q) of being observed (i.e., caught), where $0 < q < 1$.
3. Becker (1968) argues that criminals are rational and thus commit crimes when the marginal utility of crime is greater than the marginal cost. Thus, we assume that policyholders misrepresent themselves or defraud insurance companies if the marginal utility of such negative behaviors is larger than the marginal disutility.
4. Fraud and misrepresentation yield a monetary benefit, π, where $\pi > 0$, and this benefit can be manifested through falsely "earned" insurance proceeds, less expensive premiums paid than would otherwise be available, or through other financial rewards.
5. If uncovered subject to probability q, fraud and misrepresentation are punishable with a monetary penalty, μ, where μ is a negative value. This cost of being caught in the commission of risk misrepresentation or claims fraud, μ, is high; more specifically, $|\mu| > |\pi|$. This makes intuitive sense since the penalty for dishonesty μ needs necessarily to be larger than the reward for dishonesty π to create a material deterrent for an opportunistic policyholder. This penalty μ is a function of the insurance contract premium (V) and the ae amount of insurance purchased (α), with $\mu(\alpha, V)$ exogenously determined by law.
6. As introduced previously, Ω, is a subjective, nonmonetary value. Again, nonmonetary benefits can be positive (e.g., a Muslim follows Islamic laws, $+\Omega$) or negative (e.g., a Muslim violates Islamic laws, $-\Omega$). Although religious adherence (e.g., following or violating Islamic laws) impacts Ω, personal integrity and reputation do also. Since policyholder behavior is imperfectly observable (and may indeed be hidden), policyholder behavior may not hold significant

reputation value. Furthermore, unlike the model of the Takaful insurance purchase, this model assumes that nonreligious Muslims and non-Muslims place intrinsic value on their own behavior as insureds. For these reasons, Ω is more complicated in this extension of our model than in the insurance purchase model.

7. A risk-averse Takaful insurance buyer can engage in honest or opportunist behaviors. Perrin (2000) finds that religious individuals (e.g., individuals believing in life after death, attending group bible study and/or attending church) are more honest than nonreligious individuals. This is also supported by Stavrova, and Siegers (2014), who examine the relationship between an individual's religion and fraud, especially insurance fraud (e.g., claims exaggeration and the misuse of insurance), and whether such a relationship is influenced by the enforcement of religiosity across countries. The authors provide significant robust results that irreligious individuals are more likely to commit insurance fraud across countries than are religious individuals.
8. We acknowledge that regardless of religion or the level of religiosity, a policyholder may be honest (*H*) and place significant (infinite) value on personal integrity and reputation. By definition, policyholder *H* always behaves honestly. We assert that the marketplace largely comprises opportunist policyholders *O*, for whom the element of religious adherence, Ω, is critical to the honesty-dishonesty decision process. Assuming that policyholder *O* is the decision maker and following the work from Perrin (2000), and Stavrova, and Siegers (2014), we assume here that Ω holds greater value (positive or negative) for religious policyholders than for nonreligious policyholders. Thus, we accept the notion that an opportunistic policyholder, who is also religious (*R*), is more likely to behave honestly than an irreligious policyholder, although the religious policyholder may not necessarily behave perfectly honestly under all circumstances; thus, Ω(R) is increasing in R.
9. For any policyholder, the utility associated with the nonmonetary value U(Ω) may be greater or less than the utility associated with the monetary benefit of risky behaviors U(π). The insurance purchase decision is assumed to be perfectly observable and maintained over a significant period of time (generally, at least a year); therefore, we assume for the purchase decision that U(Ω) > $U(V^T\text{-}V^C)$. The risk behavior, however, is not perfectly observable and may be considered a momentary decision (and thus arguably easier for the religious insured to repent of or imagine earning an offsetting credit for).

Given these assumptions and restating the findings of Perrin (2000), Stavrova, and Siegers (2014) to indicate the Islamic influence of our study, we assert that religious Muslims are less likely to demonstrate negative policyholder behaviors than nonreligious Muslims, all else being equal. Thus, we are likely to observe less asymmetric information in the religious population than in the nonreligious population.

Religious policyholders who have purchased full insurance are likely to emphasize the nonmonetary costs and benefits, as both are large. For nonreligious policyholders, the nonmonetary costs and benefits are expected to hold a lower value in their utility function (i.e., Ω may comprise personal integrity and reputation value but not religious value). An opportunistic policyholder's expected utility maximization problem associated with the decision to behave dishonestly can be characterized by:

$$\text{Max } EU_O = qU(W - V - \mu(\alpha,V) - \Omega(R)) + (1 - q)U(W - V + \pi - \Omega(R)), \quad (8)$$

The first-order condition for maximizing (8) with respect to *R* yields

$$R^*: -q \bullet \Omega'(R) \bullet U'(W - V - \mu - \Omega(R)) - (1\text{-}q) \bullet \Omega'(R) \bullet U'(W - V + \pi - \Omega(R)) = 0 \quad (9)$$

Rearranging (9), the opportunistic policyholder is indifferent between honest and dishonest behavior when the marginal expected utility of rewards, $(1\text{-}q) \bullet \Omega'(R) \bullet U'(W - V + \pi - \Omega(R))$, equals the marginal expected disutility of costs, $q \bullet \Omega'(R) \bullet U'(W - V - \mu - \Omega(R))$.

For a religious opportunistic policyholder (i.e., R=1) to behave dishonestly, the marginal utility of π expected must be greater than the marginal disutility that is expected from μ and Ω. This opportunistic decision hinges on the sizes of q, μ, π and Ω, where q is small, subject to $0 < q < 1$, and $\mu > \pi$.

Small q, large π, small Ω. If q is sufficiently small and π sufficiently large so that the expected utility from dishonesty approaches the expected disutility of μ, an opportunistic, religious decision maker is left to focus on the marginal disutility of Ω. If Ω is small (e.g., under Sharia law, the "bad deed" debit from Allah is low; under Catholicism, the 'sin' is venial), a policyholder may behave dishonestly. This is the case in which a religious opportunist is most likely to behave like an irreligious opportunist.

Larger q, small π, large Ω. When q is inherently small, combined with a small potential reward for misbehavior and a large Ω, larger values of this small chance of being caught lead to a smaller propensity to behave dishonestly.

Small q, large π, large Ω. We assume that arguably, the most difficult behavioral decision for an opportunist who is also religious occurs when the potential reward is great relative to the potential penalty and yet Ω is also large (and negative).

2.3.4 Concluding Remarks on the Theoretical Model

The theoretical model developed here captures the role of religiosity in the Takaful insurance demand and its effect on the negative drivers of selection and asymmetric information (e.g., fraud, misrepresentation, and withheld information). In this simple model, we show that while holding other factors constant, religiosity (as an intrinsic value driver) is a key factor in Takaful insurance demand for religious Muslims and can be a significant driver in reducing negative behaviors that influence the presence of asymmetric information. More specifically, we show that the nonmonetary, intrinsic value ($\pm \Omega$) that religious Muslims would gain from God "Allah" by following the Islamic laws explain the demand for Takaful insurance products.

Religious Muslim policyholders emphasize religious value in their decisions. We assert that this intrinsic (nonmonetary) value also mitigates their negative behaviors that influence the presence and/or the level of asymmetric information.

We argue that the utility functions of irreligious and non-Muslim policyholders emphasize the price differences between Takaful and commercial insurance products, and thus, these policyholders would demand Takaful insurance products only if their coverage terms are superior to those of the commercial insurance counterparts, since the price of Takaful insurance is higher. Furthermore, we assert that the intrinsic costs of compromising personal integrity, reputation and/or religiosity have less mitigating influence on the negative economic behaviors of irreligious Muslims (Ω for irreligious Muslims < Ω for religious Muslims; no assumption made regarding Ω for non-Muslims versus religious Muslims). Therefore, irreligious Muslim policyholders are more likely to perform negative behaviors and violate Islamic laws as insurance policyholders, and the level of asymmetric information is likely to be higher among the irreligious Muslim population than among the religious Muslim population.

Although we have highlighted the influence of religiosity on Takaful insurance demand and asymmetric information, our research raises additional questions regarding the extent to which this demand and asymmetric information may change if we relax some of the assumptions (e.g., risk factors, initial wealth, and Takaful-commercial prices differential) from our simple model. For instance, what if Takaful insurance prices experience a significant reduction or increase? Furthermore, is there a relationship between Takaful insurance demand and the level of asymmetric information present? Given a policyholder's religion and other economic preferences, the rationale behind Takaful insurance demand and its potential effect on negative economic behaviors is important to understanding individual insurance decisions and is thus worth further investigation. We encourage insurance economists to further explore the roles of religion and culture on Takaful insurance demand and their effects on the prevalence of asymmetric information and fraud across all insurance markets (e.g., Takaful vs. commercial). This section serves as a cornerstone for future extensions and should help researchers to develop advanced theoretical models to better understand the role of religion on asymmetric information and Takaful insurance demand. Our efforts also resonate with the call from Schlesinger (2013), stating, "…I look forward to seeing the directions in which the theory of insurance demand is expanded in the years to come, and am encouraged to know that some of you who are reading this chapter will be playing a role in this development."

2.4 Theoretical Background and Development of the Hypotheses

As we mentioned previously, the existence of asymmetric information and adverse selection in many types of conventional insurance has been widely documented. However, there has been no focus on whether asymmetric information and/or adverse selection are present in the Takaful (Islamic) insurance market and whether religion and culture influence the extent of asymmetric information and/or adverse selection across different lines of insurance. Cohen and Siegelman (2010) provide a detailed literature review of empirical studies that focus on conventional insurance markets across different lines of business. The authors note that empirical research in this area has generated mixed results across all commercial insurance lines and that insurance economists generally agree that asymmetric information and/or selection is present in most insurance markets. Our hypotheses are developed from the broader literature on the relationship between religion, culture, and economic behavior, as well as the literature on asymmetric information and adverse selection in conventional insurance markets.

2.4.1 Asymmetric Information and Culture: Is Asymmetric Information Present in the Takaful Health Insurance Market?

The behavior of policyholders is important to insurance companies because individuals have information that may be related to their level of risk but is unknown to insurers. According to Akerlof (1970) and Rothshild and Stiglitz (1976), this information asymmetry can lead to adverse selection.[20] However, given the potential for moral hazard in insurance and the difficulty of ruling out the moral hazard, the results from our statistical tests should be interpreted as the aggregate effect of asymmetric information. Furthermore, since much of the literature considers the effects of adverse selection on insurance behavior,

[20] For instance, the theoretical work of Rothschild and Stiglitz (1976) provides evidence that profit-maximizing policymakers are incapable of distinguishing high risks from low risks in a competitive market. Thus, this would make low-risk individuals worse off due to the presence of high-risk individuals, as it would motivate them to buy less insurance than they would in markets where asymmetric information is not present.

we will discuss this literature here. Cummins, Smith, Vance, and VanDerhei (1983) define adverse selection as "… the tendency of high-risks to purchase insurance or to purchase more insurance coverage than do low-risks."

Researchers have empirically examined the presence of adverse selection by testing the relationship between insurance demand and insurance utilization (measured by claim behavior) after the seminal contributions of Akerlof (1970) and Rothshild and Stiglitz (1976). Significant research in this context provides evidence that adverse selection is present in the healthcare insurance markets (e.g., Panhans 2019, Cutler and Zeckhauser 2000, Simon 2005, Eling, Jia, and Yao 2017), while other studies reveal mixed results, including those by Farley and Monheit (1985), Wrightson, Genuardi, and Stephens (1987), and Long, Settle, Wrightson (1988), and Panhans (2019).

Given the lack of detailed data on the policyholders' religious affiliation and nationality, prior studies could not examine the roles of religion and culture as potential influential factors of asymmetric information in the unique "Takaful" health insurance market that adheres to the principles and teachings of Islam. The only study that we are aware of on this "Takaful" health insurance market is the study by Rahman and Daud (2010). In their 2-year survey in the Malaysian insurance industry, they find no evidence of moral hazard in the "Takaful" health insurance market. However, when considering nationality, race and religion, they argue that tendencies toward making exaggerated claims, engaging in negative behavioral patterns and withholding information are relatively low among Muslims. Their survey was very limited, as it was conducted exclusively on Malaysian respondents for one year, and 95 percent of their respondents were Muslims; hence, there was not much variation in religious groups. Given this limitation, our work adds to the existing literature by studying asymmetric information in the Takaful insurance market by more effectively using a large dataset that contains information about religious affiliation and nationality for every policyholder.

Regarding the empirical work that tests the existence of asymmetric information in the healthcare insurance market, after the seminal work of Chiappori and Salanie (2000),[21] it has been a common practice to examine whether there is a positive relationship between claim behavior and insurance coverage characteristics. However, the choices of using proxies for insurance coverage have not been settled. For instance, subsequent researchers have utilized the deductible choice as a proxy for coverage

[21] For more details on the models used to test selection and asymmetric information in the insurance literature, see Cohen and Siegelman (2010).

(e.g., Puelz and Snow, 1994 and Cohen and Einav, 2007). However, Born and Sirmons (2019) argue that the proxies for risk depend on the ex ante and ex post measures.[22]

Given the theoretical and empirical work on selection and asymmetric information, a potential information disadvantage occurs when an insurance company is not able to observe full information about its policyholders. If insurers can observe a policyholder's past claim records, the information disadvantage will be reduced, and they will control the problem of adverse selection more effectively. However, the Insurance Research Council (1991) states that a new policyholder's self-reporting of past claims is incomplete or inaccurate, and research finds that insurers cannot observe a new policyholder's past claim records (Cohen, 2005). This finding suggests that insurers are more likely to experience an information disadvantage with new customers. Although mixed results were reported by Cohen and Siegelman (2010) across all insurance markets, insurance economists generally agree that asymmetric information and/or selection exist in most insurance markets. This argument applies to both commercial and *Takaful* insurance markets. As such, we posit the following hypothesis.

H1. Asymmetric information is present in the Takaful healthcare insurance market.

The key question is whether asymmetric information is present in the Takaful healthcare insurance market, given that all the drivers of asymmetric information—misrepresentation, fraud, withheld information, and lying—are rejected in all forms according to Islamic laws. Do such Islamic laws influence the current and potential policyholders' negative behaviors? If so, to what extent do these laws influence the level of asymmetric information in the *Takaful* insurance market? These questions motivate us to discuss the potential effect of religion and Islamic culture on asymmetric information in the next section.

2.4.2 Asymmetric Information and Religion: Does Religion Influence Economic Behavior?

After the seminal work of Adam Smith (1776) and Weber (1905), economists have examined the effect of religion on economic behavior. However, Iyer (2016) states that the seminal work of Iannaccone (1998) introduces to a large audience of economists the economics of religion—as a new field of study—in which there has been a substantial increase in the number of published papers on this research area,

[22] Born and Sirmons (2019) state that while most researchers utilize ex post risk measures (e.g., the number of claims and the average claims amount), some researchers utilize measures of long-term conditions (e.g., Doiron, Jones and Savage, 2008) and measures of self-assessed health (Browne and Doerphinghaus, 1993) as proxies for ex ante risks. The authors state that ex ante risk measures capture adverse selection but not moral hazard, as such ex ante risks are measured prior the policy's effective date. However, the ex post risk-coverage correlation reflects both moral hazard and adverse selection.

and an increasingly large number of international scholars have shown interest by connecting religion to different fields of economics[23].

Economists have examined the economic effect of religion at the micro and macro levels. For instance, at the micro level, researchers have examined religious affiliations and their effects on making individual-level decisions with regard to divorce and marriage (e.g., Heaton and Pratt, 1990 and Lehrer and Chiswick, 1993), consuming alcohol and drugs (e.g., Cochran and Akers, 1989), committing crimes (e.g., Evans et al, 1995), gambling (e.g., Diaz, 2000) and committing suicide (e.g., Bainbridge, 1989). At the macro level, researchers have answered questions regarding the influence of religion on economic growth (e.g., Hilary and Hui 2009) and the effect of religion on economic attitudes (e.g., Guiso, Sapienza, and Zingales 2003)[24].

The economics of religion have even expanded to areas outside the economic literature, in which within the finance literature, researchers have examined the effect of religion on companies' and investors' behaviors (e.g., Hilary and Hui 2009)[25] and on people's religiosity and risk aversion (e.g., Millerand Hoffmann, 1995 and Bachman, Johnston, and O'Malley 1993). Within the insurance literature, which is the focus of our study, insurance economists have examined how religion, as a proxy for national culture, influences insurance demand across various countries, including some Muslim countries. Insurance economists state that while policyholders are assumed to make rational decisions, they may respond to insurance demand according to their cultural perceptions and religious beliefs (e.g., Outreville 2018, Trinh, Nguyen, and Sgro 2016, and Park and Lemaire 2012). One more explanation, as noted in Park and Lemaire (2012), is that "… some religious people, especially Muslims, believe that reliance on insurance to protect one's life or property results from distrust in God's protective care". Demand for insurance is a function of a person's lifetime allocation and utility maximization which depend on wealth, income, risk aversion and the cost of insurance. Risk aversion is a function of personal demographics which include culture and religion (Yaari 1965). Risk aversion is reported to be positively associated with religiosity (e.g., Hilary and Hui 2009, and Noussair et al 2013). Miller and Hoffmann (1995) find level

[23] Iyer (2016) states that after the seminal work of Iannaccone (1998), "…The economics of religion has made important strides, with studies that now encompass economic theory, public economics, experimental economics, macroeconomics of growth, economic history, and economic development." Further, the economics of religion has its own annual meeting and its own JEL classification number (Z12).

[24] For more details about the effect of religion and culture on economic behavior/outcomes in the economics literature, see Iannaccone (1998), Iyer (2016), Guiso, Sapienza and Zingales (2003) and Guiso, Paola and Zingales (2009).

[25] Hilary and Hui (2009) evaluate the role of corporate culture on the companies' behavior. The authors report that the degree of risk exposure is significantly lower for firms located in highly religious countries than for firms in other countries. For more details on the effect of religion on companies' and investors' behavior in the finance literature, see Hilary and Hui (2009).

of individual religiosity is consistent with being more risk averse. This is supported by Osoba (2003) who finds that risk averse individuals attend church more often than risk seeking individuals.

This insurance literature clearly suggests that religious Muslim policyholders may behave differently in terms of buying insurance and that their negative claiming behavior is different from that of other policyholders due to their Islamic beliefs. Furthermore, as mentioned previously, according to *Islamic* law, *all drivers of asymmetric information, namely, misrepresentation, fraud, withheld information, and lying, are rejected in all forms.* Given the existing literature on how religion influences economic behavior, especially insurance demand, the extent of asymmetric information may vary across religious groups. Furthermore, given that the Takaful product is an *Islamic law*-compliant insurance product, it is more expensive than other commercial insurance products (Coolen-Maturi 2013), and it is the only alternative to conventional insurance (Kwon, 2007). This phenomenon may suggest that religious Muslims care more about the religious value of the product and respond to demanding Takaful insurance according to their religious beliefs and cultural perceptions, regardless of how expensive Takaful coverage is[26]. Therefore, we expect the Muslim proportion of our sample to be more religious than the non-Muslim proportion. In addition, we expect the relatively religious to be less inclined to engage in negative behaviors that would present asymmetric information. Thus, we formally hypothesize the following.

H2. The extent of asymmetric information is lower for Muslim policyholders than for non-Muslim policyholders.

2.4.3 Asymmetric Information and Culture: Do Hofstede's Cultural Dimensions Influence Asymmetric Information?

Next, we consider the effect of Hofstede's cultural dimensions on asymmetric information. Given the existing literature on the influence of Hofstede’s cultural variables on economic behavior, particularly the insurance demand literature, and on the overlap of this literature with the literature on religion, we expect to find that asymmetric information is associated with Hofstede’s cultural variables. However, among Hofstede’s (1984) cultural dimensions, insurance economists have mainly utilized four of these variables to explain the demand for insurance, namely, power distance, individualism, masculinity, and uncertainty avoidance (e.g., Park and Lemaire 2012). In this study, we follow their lead and their findings

[26] Coolen-Maturi (2013) shows that moderate Muslims in the UK are willing to demand Takaful insurance products only if their prices and coverage are as competitive as conventional ones, which suggests that moderate Muslims care about the economic value of insurance products and are less religious than Muslims who demand Takaful insurance products.

to explain the potential association between the four dimensions of Hofstede's cultural variables and the level of asymmetric information/adverse selection in the *Takaful* healthcare insurance market.

Power Distance (PDI) measures the degree of power inequality among people (e.g., privileges, education, power, and wealth). Chui and Kwok (2008) argue that a high power distance score reduces the demand for insurance. Trinh, Nguyen, and Sgro (2016) report consistent results, in which the higher the degree of power distance is, the lower the insurance expenditure. According to the above findings, *we expect a negative relationship between a high score power distance score and the level of asymmetric information/adverse selection. Thus, we formally hypothesize the following.*

H3. The extent of asymmetric information/adverse selection is negatively associated with a high power distance score.

Individualism (IDV) measures the extent to which people prefer to act as individuals or as members of a group. Trinh, Nguyen, and Sgro (2016) and Chui and Kwok (2008) argue that relying on insurance among individualistic countries is the norm and find a positive relationship between insurance consumption and a high level of individualism. *Given these findings, we expect the level of adverse selection/asymmetric information to be positively associated with individualism, especially among developed countries. Thus, we formally posit the following.*

H3B. The extent of asymmetric information/adverse selection is positively associated with a high individualism score.

Masculinity (MAS). Park and Lemaire (2012) state that "…masculinity evaluates whether biological gender differences impact roles in social activities." Chui and Kwok (2008) report that feminine societies demand more life insurance than masculine societies, as feminine societies are more protective against financial risk and care more about the needs of their families. Furthermore, the author reports an ambiguous relationship between masculinity/femininity and non-life insurance demand. On the other hand, Trinh, Nguyen, and Sgro (2016) argue that the lower the masculinity score is (in feminine societies), the larger the population's non-life insurance consumption. *Given these conflicting findings, we expect the effect of masculinity on the level of adverse selection/asymmetric information to be ambiguous. Thus, we formally postulate the following.*

H3C. The extent of adverse selection/asymmetric information is associated with a high (low) masculinity score.

Uncertainty avoidance (UAI) measures an individual's tolerance to ambiguous situations. Hofstede (2001) reveals that risk-averse individuals tend to have a high UAI, which compels them to avoid uncertain events and unnecessary risks. As one would expect, Park and Lemaire (2012) and Trinh, Nguyen, and Sgro (2016) report that a higher score of uncertainty avoidance is positively associated with larger non-life insurance consumption. Most importantly, Park and Lemaire (2012) reveal that insurance economists' proxy for risk aversion is correlated with uncertainty avoidance. Interestingly, risk aversion is reported to be positively associated with religiosity (e.g., Hilary and Hui 2009, and Noussair et al 2013). This report suggests that individuals with a high UAI are likely to be religious and risk averse. Further, *this may suggest that in regard to insurance utilization, individuals with a high UAI are less likely to misuse their insurance, defraud insurance companies, withhold information, and misrepresent themselves. All these negative behaviors are major contributors to adverse selection/asymmetric information. Thus, we formally postulate the following.*

H3. The extent of asymmetric information/adverse selection is positively associated with a low uncertainty avoidance score.

2.5 Data and Methodology

2.5.1 Disjoint Between the Theoretical Model and Empirical Work

To test the mitigation effect of Muslim religiosity on the asymmetric information developed in our theoretical model (see Section 2.3), we should ideally compare religious Muslims with nonreligious Muslims. However, as shown in other studies, it is extremely difficult to measure the extent of religiosity of Muslim groups, which explains the lack of data on Muslim religiosity. Fortunately, we have a proprietary dataset from the United Arab Emirates (UAE) that provides the religious affiliation for each policyholder (i.e., Muslim, Christian, atheist, etc.).

We utilize the beneficiary's reported religious preference as Muslim or not Muslim and compare the Muslim (61% of the sample) versus non-Muslim (39%) individuals in our empirical model. Table 2.3 summarizes an *ideal* case of a comparison between religious Muslims (Group A) and nonreligious Muslims (Group B) and our approach of comparing Muslims (Group A ≈ Groups A+B) and non-Muslims (Group B ≈ Groups C+D). Below, we argue that our approach is reasonable for the following four reasons:

Table 2.3. The Potential Proportion of Irreligious Muslims in the Sample

Columns	(1)	(2)
	Muslim Sample	Non-Muslim Sample
Religious	A	C
Irreligious	B	D
Total	40,484 policyholders (61 percent)	25,856 policyholders (39 percent)

First, given that Takaful insurance is compliant with Sharia law and more expensive than conventional insurance, it is quite reasonable to assume that the majority of Muslims in the Muslim category (Groups A+B) are religious Muslims (Group A)[27].

Second, irreligious Muslims (Group B) are likely to represent only a random and small percentage of the Muslim category. The misclassification through measurement error of irreligious Muslims (Group B) as religious Muslims (Group A) is also likely to work against us in finding any significant results regarding the behavior of religious Muslims.

Third, within the non-Muslim category (Groups C+D), there are religious non-Muslims (Group C) and irreligious non-Muslims (Group D). However, we cannot obtain the extent of religiosity of the non-Muslim group to separate Group D from Group C, except for a small percentage of irreligious non-Muslims who reported their religious affiliation as being atheist, agnostic or having no religion (Group C).

Fourth, we argue that in our Takaful insurance sample, Group C is likely to capture only a small and random proportion of the non-Muslim sample as religious; therefore, the Group 3 results are unlikely to influence the results within the non-Muslim population, especially those results regarding the behavior of irreligious non-Muslim policyholders (Group D). As we mentioned previously, the rationale behind this argument is that (observing) Christians, for instance, break no Christian law by purchasing Takaful

[27] The rationale behind this argument is explained within our theoretical model (Section 3), in which we show that the nonmonetary, intrinsic value ($\pm \Omega$) that religious Muslims would gain from God ("Allah") by following the Islamic laws explain the demand for Takaful insurance products and can be a significant driver of the reduction in the negative behaviors that influence the presence of asymmetric information.

insurance. Indeed, a religious Christian might find Takaful insurance to be superior to commercial options available because of the principles by which it operates.

2.5.2 Data

To test our hypotheses, we use a proprietary dataset that contains all policyholders and individual medical claims from a UAE Takaful insurance company from 2014 to 2015. These data are useful for considering the association between religion, culture and asymmetric information for the following *three* reasons. First, this dataset contains policyholders who are from 141 different countries but who all reside in the UAE (see Appendix B). The insurance company also provided us with information about the individual policyholder's religious preference *as reported by the policyholder*, which enables us to examine the impact of religion on claim performance (H2). Furthermore, due to the availability of the nationality of each policyholder in this dataset, we can assign cultural variables based on nationality to investigate how each of the cultural measures influences asymmetric information (H3A, B, C, and D). Second, our dataset contains information that the insurer uses to make underwriting and pricing decisions, such as health and family status. After controlling for these pricing factors, this study is in a strong position to estimate the impacts of religion and culture on claim performance. Third, this dataset contains the characteristics of each policyholder's insurance policy (i.e., policy commencement and termination dates) and the medical claim characteristics for each insurance claim made (i.e., filing date). With these various data available to us, we can create a panel of daily insurance claim behavior for the enrolled policyholders.

The variables utilized in this paper are listed in Table 2.4 alongside brief descriptions. To consider the association between risk and insurance coverage, we use the following four measures of claim performance as a measure of risk: (1) the filing of a claim by the policyholder; (2) the payment of a claim to the policyholder; (3) the amount claimed by the policyholder, and (4) the amount paid to the policyholder. To proxy for insurance coverage, we take the natural logarithm of the annual maximum level of insurance coverage. We include the following five control variables: (1) gender; (2) age; (3) relationship to the principal policyholder (i.e., spouse or children); (4) marital status; (5) educational background, and (6) policy length (in days) for each beneficiary under an insurance policy.

To identify religiosity, we use the beneficiary's reported religious preference as being Muslim or not. Since the insurance product is compliant with Islamic law, we assume that the Muslims in the dataset are likely to be relatively more religious than the non-Muslims who purchase Takaful insurance. Finally, we assign the Hofstede cultural variable by using the reported nationality of the policyholders. Thus, we assume that any given policyholder is representative of his/her country's culture, as measured by the cultural indices utilized in this paper. Our assumption is likely to hold in our sample, given that the majority of the policyholders (66 percent) in our sample are immigrants from other countries. Liu (2016) finds that "*…when individuals emigrate from their native country to a new country, their cultural beliefs and values travel with them, but their external economic and institutional environment is left behind. Moreover, immigrants not only bring their beliefs and values to the new country, they also pass down these beliefs to their descendants.*" Thus, the policyholders' nationality and their claiming behavior can be used as a proxy for culture for not only their country but also their descendants.

For our dataset, Table 2.5 reports the summary statistics between the two subsample groups, namely, the Muslim (*religious*) and non-Muslim samples. These two samples are generally quite different, especially the Muslim sample, which seems to have more males, be in better health, be younger, have fewer married individuals and have more children. Furthermore, the difference in claim behavior is largely suggestive of what we would expect in Hypothesis 2; i.e., the religious individuals are less likely to engage in behavior representative of information asymmetry. Given the differences in these two groups, it will be important to use the previous demographic measures as controls.

Table 2.4. Descriptions of the Variables

Variable	Description
Dependent variables:	
Claim Filed	1 if any claim was made by beneficiary
Claim Paid	1 if any claim was paid to beneficiary
Claim Amount Filed	Claim amount made by the beneficiary
Claim Amount Paid	Claim amount paid to the beneficiary
Key variables of interest	
Coverage Amount	Logged maximum amount of insurance coverage
Muslim	1 if the policyholder reported being Muslim
Religious Holiday	1 if a day is one of the religious Muslim holidays (e.g., Ramadan, Eid Al-Fitr following Ramadan, and Eid Al-Adha)
Muslim* Religious Holiday	Interaction term between Muslims and religious holidays
Friday	1 if the policyholder visits a medical facility on Fridays (Friday effect) and 0 otherwise
Muslim* Friday	Interaction term between Muslims and “religious” Fridays
Sunday	1 if the policyholder visits a medical facility on Sundays (Sunday effect) and 0 otherwise
Muslim*Sunday	Interaction term between Muslims and Sunday
Control Variables	
Policy Length	Length of policy in months
Age	Age of beneficiary
Male	1 if male
Married	1 if married
Principal	1 if the principal of the policy
Spouse	1 if the spouse of the policyholder
Child Dependent	1 if the child dependent of the policyholder
College Degree	1 if the policyholder is educated (college degree)
Chronic	1 if the claim made on a chronic condition
Moderators:	
Individualism (IDV)	Measures the extent to which people prefer to act as individuals or as members of a group
Masculinity (MAS)	“…evaluates whether biological gender differences impact roles in social activities.”
Uncertainty Avoidance (UAI)	Measures the individual’s tolerance to ambiguous situations
Power Distance (PDI)	Measures the degree of power inequality among people (e.g., privileges, education, power, and wealth)

Table 2.5. Muslim and Non-Muslim Summary Statistics

The table shows the summary statistics for the Muslim and non-Muslim populations. Columns 1 and 2 report the mean values and standard deviation for the Muslim subsample, respectively. Columns 3 and 4 report the mean values and standard deviation for the non-Muslim subsample, respectively. All variable definitions are shown in Table 2-4. The significant difference (p-values) between the two samples is reported in Column 6.

Column	(1)	(2)	(3)	(4)	(5)=(1)-(3)	(6)
	Muslim (N=40,484 obs)		Non-Muslim (N=25,856 obs)			
	Mean	SD	Mean	SD	Difference	p-values
Claims Made	0.6311	(0.483)	0.6515	(0.477)	-0.0204	0.000
Claims Paid	0.6296	(0.483)	0.6500	(0.477)	-0.0204	0.000
Average Claim Filed	185.5118	(1612.474)	312.5165	(2015.441)	-127.0047	0.000
Average Claim Paid	168.1608	(1593.904)	277.2061	(1910.188)	-109.0453	0.000
Coverage Amount	12.6754	(0.246)	12.7008	(0.258)	-0.0254	0.000
Policy Length	500.1877	(202.544)	502.6551	(202.921)	-2.4674	0.126
Age	28.6584	(13.962)	32.6378	(13.742)	-3.9794	0.000
Male	0.7710	(0.420)	0.6991	(0.459)	0.0719	0.000
Married	0.6180	(0.486)	0.6294	(0.483)	-0.0114	0.003
Principal	0.7098	(0.454)	0.7640	(0.425)	-0.0542	0.000
Spouse	0.1010	(0.301)	0.1149	(0.319)	-0.0139	0.000
Child Dependent	0.1893	(0.392)	0.1212	(0.326)	0.0681	0.000
Chronic	0.0671	(0.25)	0.1104	(0.313)	-0.0433	0.000
Education (College Degree)	0.2335	(0.423)	0.2343	(0.424)	-0.0008	0.805
Individualism	36.1644	(13.232)	40.5887	(15.166)	-4.4243	0.000
Masculinity	53.2565	(4.458)	52.6528	(12.631)	0.6037	0.000
Uncertainty Avoidance Index	54.1883	(15.938)	47.2567	(13.568)	6.9316	0.000
Power Distance Index	75.1077	(8.912)	75.3717	(15.157)	-0.2640	0.008

2.5.3 Methodology

To examine the relationship between an individual's religious and cultural backgrounds and asymmetric information, we employ an empirical strategy similar to that used in Finkelstein and Poterba (2004) and estimate a linear relationship between insurance coverage and claim behavior, controlling for individual observables. More specifically, we estimate the following linear models for individual i on day t:

$$Claim_{it} = \alpha + \Omega lnCovAmt_i + \theta X_{it} + \beta R_i + \varepsilon_{it} \tag{2.1}$$

We define *Claim* as one of the four claim variables discussed in the previous section. Next, *lnCovAmt* is the logged yearly maximum insurance coverage level. We include *X as* a vector of individual-specific observable characteristics. Included in *X* are gender, age, marital status, policy length, relationship to the principal of the policy, education, chronic condition, and nationality. The coefficient of interest for this regression is β. When considering the relationship between religiosity and asymmetric information, we define R as a binary variable denoting Muslim as the policyholder's reported religious preference. When considering the relationship between culture and asymmetric information, we define R as one of the cultural variables described in the previous section. According to Equation 2.1, if asymmetric information is present, we expect $\Omega>0$. Note that Ω has no causal interpretation due to the nature of the problem of asymmetric information. Chiappori and Salanie (2000) report that a positive correlation between claim performance variables and insurance coverage suggests that high-risk individuals buy more generous coverage. This positive correlation indicates the possibility of asymmetric information. An insignificant correlation between risk and coverage may suggest no asymmetric information.

Additionally, to report more robust results, we follow Finkelstein and McGarry (2006) and use the following two reduced models to examine the presence of asymmetric information:

$$Claim_{it} = \alpha + \theta X_{it} + \beta R_i + \varepsilon_{it} \tag{2.2}$$

$$lnCovAmt_{it} = \alpha + \theta X_{it} + \beta R_i + \gamma_{it} \tag{2.3}$$

According to the models above, the positive correlation between the residuals ϵ and γ is a necessary condition for the presence of adverse selection or asymmetric information, which suggests that the relationship between coverage choice and claim occurrence is dependent, after controlling for observables. An insignificant correlation between ϵ and γ suggests that there is no adverse selection or asymmetric information. Thus, following Finkelstein and McGarry (2006) and Chiappori and Salanie (2000), we will use the classical model as our core model and the reduced form as a robustness test.

2.5.4 Additional Analysis and Robust Results

To show that our main results—the negative relationship between the level of religiosity and negative claiming behaviors—are robust, we conduct the following three additional tests.

First, we consider the effect of major religious holidays and/or events for Muslims (e.g., the holy month of Ramadan) by incorporating a separate religious holiday binary variable and then interacting the Muslim variable with our religious Muslim holiday binary variable. We do this because religious holidays provide good exogenous variation in dates because they are based on the Islamic lunar calendar and change across years. For instance, Campante and Yanagizawa-Drott (2015) explain that the solar year is 10 to 11 days longer than the Islamic lunar year, which causes the lunar months to rotate over the years in a cycle of 33 years. This process causes the number of fasting hours during Ramadan to vary depending on the month/year in which Ramadan occurs[28]. Thus, Muslim religious events/holidays provide us with ideal exogenous variation to examine the relationship between religiosity and the policyholder's claiming behavior.

Furthermore, religious events/holidays have been shown to have a significant effect on economic growth and stock markets. For instance—given a country's latitude and the rotation of the Islamic "lunar" calendar—Campante and Yanagizawa-Drott (2015) examine the effect of fasting during Ramadan on economic growth and happiness, that is, "subjective well-being". The authors report that fasting reduces an individual's productivity and thus economic growth. Furthermore, Ramadan has a positive impact on an individual's subjective well-being. Furthermore, researchers have examined the effect of Islamic religious holidays (e.g., Hajj effect) on the Islamic stock markets[29] (e.g., Wasiuzzaman, 2017), the effects of seasonality and Hajj pilgrimage on the Saudi stock market (e.g., Wasiuzzaman, 2018), and the impact of Ramadan on herding behavior in the Pakistani Stock Market (e.g., Yousaf, Ali, and Shah 2018).

Given this literature, we argue that people may be more religious and spiritual during religious events and tend to not perform any bad acts that could benefit them and/or harm anybody else (e.g., negative claim behavior, lying, misrepresentation, misuse of insurance, fraud, and the withholding of information). Our argument is also supported by Campante and Yanagizawa-Drott (2015), who state that

[28] For instance, in 2018, Ramadan began on May 16 and ended on June 14. However, in 2019, Ramadan began approximately 11 days earlier, on May 6. Further, note also that the number of fasting hours differ across countries depending on the latitude. For more information, see Campante and Yanagizawa-Drott (2015). However, this study is not affected by a country's latitude, since all policyholders across 141 countries live in the UAE.

[29] The effect of non-Muslim religious holidays (e.g., Easter) on earning announcements (e.g., Pantzalis and Ucar 2014) and the relation between Lent and Advent in the Catholic Church calendar and stock price behavior (e.g., Dumitriu and Stefanescu 2017) have also been examined.

Ramadan influences the individuals' social lives and lifestyles, as they become more involved in spiritual events, such as reading the Holy Qur'an and praying at mosques.[30]

Second, we consider the most important "religious" day of the week, Friday, as the "Jumu'ah" effect". In fact, we control for the presence of the Friday effect on claiming behavior and asymmetric information. According to prophet Muhammad's teaching, "…Friday is the best of days, and the most virtuous in the sight of Allah. Further, the Muslim who passes away on the night or during the day of Friday, Allah saves him from the punishment of the grave". In the Islamic world, Muslims are often more spiritual, religious, and forgiving on Fridays. We argue that such holy days may reduce abusive behaviors (e.g., fraud and negative claiming behavior) among Muslims. Thus, we argue that on Fridays, Muslims utilize their medical insurance only for medical emergencies or due to availability/selection issues (e.g., full-time jobs), while individuals who work full time during the weekdays and suffer from illness may decide to wait until the weekend to visit medical facilities. Therefore, this phenomenon may reduce the number of medical visits and claims by Muslims on Fridays but may increase the size of their claims (claims amount and settlement).

Third, we control for the "Sunday effect" in the insurance market and examine whether it is more/less pronounced in Muslim claiming behavior than in non-Muslim claiming behavior. This effect is particularly distinct in the disability, medical and workers' compensation insurance markets. Campolieti and Hyatt (2006) state that employees are likely to take advantage of workers' compensation insurance and file more claims on Mondays (after the weekend) than on other days. Insurance economists have become interested in examining and justifying the rationale for this phenomenon. For instance, Campolieti and Hyatt (2006) report two possible explanations. The first explanation is the psychological effect, in which after a long rest over the weekend, employees become more susceptible to any injuries when they go back to work on Sundays. The second explanation is the moral hazard effect, in which workers' compensation insurance covers the cost of any medical treatments and loss of income due to any work-related diseases or injuries, especially back injuries and sprains or strains. This phenomenon could justify the Sunday effect, as workers may become injured over the weekend and file insurance claims on Sunday, as workers who injure themselves on days off will not be compensated.[31]

[30] Wasiuzzaman (2018) states that religious holidays make people more devoted to God ("Allah"), enhance their faith, and make them more spiritual according to Prophet Muhammad's teaching: "… It teaches the values of mercy, forgiveness and symbolizes humility. It also helps to integrate the Muslim world, leading to a strengthening of global Islamic beliefs…"

[31] For more information about the Sunday effect, see Card and McCall (1996), Ruser (1998), and Butler, Gardner, and Kleinman (2013).

Stavrova and Siegers (2014) examine the relationship between an individual's religion, unethical behaviors, and fraud, especially insurance fraud (e.g., claims exaggeration and the misuse of insurance), and whether this relationship is influenced by the enforcement of religiosity across countries. The authors provide significant and robust results that irreligious individuals are more likely to commit insurance fraud across countries than religious individuals. Given this result, we argue that this negative claiming behavior is significantly lower among the religious population on Sundays than on other days and that the "Sunday effect" is less prevalent among the Muslim population sample.

2.6 Results

Table 2.6 presents the OLS estimates of the relationship between our measures of claims behavior (ex post risk measures) and the logged maximum insurance coverage. Consistent with our Hypothesis 1, we find a positive correlation between insurance coverage and four measures of claim behaviors, suggesting that there is significant evidence of asymmetric information in the Takaful health insurance market. Considering these results and the results for our key variable (Muslim), we find that there is asymmetric information present in the market for Takaful health insurance and that religious individuals are found to engage in less individually advantageous behavior than nonreligious individuals. More specifically, our estimates suggest a consistently negative relationship between being religious and insurance claim behavior. Furthermore, the results suggest that religious people opt for less insurance coverage than do nonreligious people. Considering the sample mean of 0.638 for claims made by individuals (per day), the coefficient on claims made suggests that religious individuals are less likely to file a claim than nonreligious individuals by 0.6% per day. If we consider the difference in the claim amounts filed, we find that religious individuals file claims that are approximately 1.2% smaller per day than those of the nonreligious individuals. Both results are similar when considering the amount actually paid by the insurance company. Our expectations are also supported by the interaction term between Muslim and religious holidays, in which religious individuals are found to engage in less individually advantageous behavior during religious holidays (e.g., Ramadan and Eid-Al-fiter) than on other days. [32]

However, one might argue that our cross-sectional results can also be explained by a medical supply shortage during the Muslim public religious holidays (i.e., Eid Al-Fitr following Ramadan and Eid Al-

[32] Note that we further utilize the seemingly unrelated regressions (SUR) model, and then, following Finkelstein and McGarry (2006), we use the bivariate model to examine the presence of asymmetric information. We generate results that are consistent with our expectation, and our key variables (Muslim, and religious holidays) are significant at the one percent level. Our Breusch-Pagan test of independent results suggest that the errors are correlated. However, the correlation among residuals is generally very small (0% or 18.73%). Therefore, the SUR results are not presented in this paper. However, they are available upon request.

Adha following the Hajj pilgrimage) and that policyholders are aware of this issue, avoiding visiting medial facilities during these religious holidays.[33] To rule out the potential effect of medical supply shortages on policyholders' claiming behaviors, we repeat our analysis by excluding Muslim public religious holidays from our interaction term between Muslim and religious holidays. The new results remain consistent with our cross-sectional results, in which religious individuals are found to engage in less individually advantageous behavior during religious holidays, hence ruling out the medical supply shortage explanation.[34]

Most of our control variables are significant at the 1 percent level, and their predicted directions are consistent with the literature and with our expectations. For example, as one would expect, having a chronic condition and being older are significantly associated with higher levels of insurance usage[35]. Moreover, consistent with Eling, Jia, and Yao (2015), spouses, children, and married people tend to utilize more insurance and claim higher amounts of insurance. Interestingly, educated Muslims claim higher amounts of insurance, and the settled amount paid to them is higher and significant at the 1 percent level. Finally, males file fewer claims than females, and there is significant evidence of a difference between genders in the amount of insurance utilization, amount filed, and amount paid.

Table 2.7 presents the coefficients of interest for the OLS regression results on the relationship between our measures of culture and the logged maximum insurance coverage. The results show the extent to which Hofstede's cultural variables influence asymmetric information given the policyholder's claim performance. The estimates for the cultural variables are significant; however, many of the estimates appear to be economically insignificant. As predicted, PDI is negatively associated with claim behavior; however, the relationship is economically small. In addition, both IDV and UAI are significantly associated with claims behavior in the direction predicted; however, these measures also seem to have an economically insignificant relationship with the likelihood of a claim being filed. According to our

---

[33] Medical facilities still operate and provide medical services on public religious holidays. However, the number of health professionals working on those days tends to be low, which may cause difficulties and delays for healthcare providers and patients. Furthermore, a patient may believe that medical staff who work on the holidays have less experience and seniority and, thus, that the provided medical services are likely to be poor. This phenomenon, in fact, may influence a patient's decision to visit medical facilities during holidays, except in the case of medical emergencies.

[34] The results are available upon request.

[35] Our summary statistics in Table 4 show that the Muslim sample seems to be younger and have better health than the non-Muslim sample. One might argue that these characteristics of the Muslim sample could be the major driver of our observation of less asymmetric information among the Muslim sample and that our religiosity measure is just a proxy for these Muslim characteristics. To rule out this alternative explanation and control for policyholders' age, we follow Finkelstein and McGarry (2006) and repeat the "primary" analysis by incorporating a separate age binary variable and dividing our sample into young versus old subsamples. The results are still consistent with our expectations in which asymmetric information is still present, and being older is significantly associated with high levels of insurance usage.

hypothesis, MAS has an ambiguous relationship with claims behavior, and we find that there is a significant negative relationship. Our control variables across all Hofstede's cultural variables are significant at the 1 percent level and consistent with the results reported in Table 2.6.

Table 2.8 shows the OLS regression results of the logged maximum insurance coverage regression on claim performance variables, where the interaction terms between Muslim and Fridays and Muslim and Sundays are the key variables (Friday*Muslim and Sunday*Muslim, respectively) to capture the effect of Fridays and Sundays on an individual's claiming behavior. In other words, based on the policyholder's claim performance, this table shows the extent to which Fridays and Sundays influence asymmetric information. Surprisingly, the results show that there is a positive significant interactive relationship between being religious and the "religious" day of the week (Friday), suggesting that there is significant asymmetric information in the Takaful health insurance market and that religious individuals file more claims than nonreligious individuals on Fridays. The results may appear to be inconsistent with our expectations. However, a potential explanation for this confounded result is the time availability and selection issue, according to which, due to time constraints (e.g., having a full-time job), some individuals who suffer from illness during the weekdays may decide to wait until the weekend to visit medical facilities.[36] This phenomenon, , may cause such policyholders to file more claims at larger amounts during the weekends and visit medical doctors during the weekdays only if they suffer from serious illness (emergencies). Thus, the Friday effect on claiming behavior is present and is more pronounced among the Muslim population than among the non-Muslim population.

Table 2.8 also shows the Sunday effect on the policyholder's claiming behavior; in the demonstration of this effect, the interaction term between the Muslim and Sunday binary variables is a key variable (Sunday*Muslim). The results show that the Sunday effect is positive and distinct among all claim behavior variables. However, when we explore the interaction term between the Muslim and Sunday binary variables, we observe a negative and significant interactive relation, suggesting that there is significant asymmetric information in the Takaful health insurance market and that religious individuals engage in less individually negative claiming behavior on Sundays. The results are consistent with our expectation that religious workers are less likely to harm themselves over the weekend and file claims

[36] According to the UAE Labor Law, the weekend runs from Friday to Saturday for most public services and the government. However, this is not the case for the private sector, for which the weekend runs from Saturday to Sunday. Furthermore, the normal working hours in the private sector are 48 hours per week or (8-10 hours per day). The working hours are reduced by two hours during Ramadan, and Friday is supposed to be a rest day for most workers each week, especially for Muslims. For more information about the working days/hours in the UAE, check the UAE Labor Law www.government.ae (Articles 65, 70, and 71 of UAE Labor Law).

on Sunday to benefit from insurance. Thus, religious individuals (workers) are less likely than nonreligious individuals to purposely become more susceptible to any injuries when they return to work on Sundays.

This finding also may suggest that moral hazard issues in medical insurance are similar to those in workers' compensation insurance, where workers may injure themselves and/or visit medical facilities during workday hours to take a workday off. If we compare the positive Friday effect among religious individuals to the negative Sunday effect among the same group, the results support our argument that a negative relationship is expected between the levels of religiosity and negative claiming behaviors, as Sunday is a workday in the UAE, and most individuals are required to work at least 8 hours. The fact that we observe a positive Friday effect in which religious individuals file more claims than nonreligious individuals on Fridays may support our earlier argument that workers with illnesses visit medical facilities on Fridays due to time constraints. Given this confounding result, more research should be conducted to further deepen our understanding of how Sunday and Friday effects may influence claiming behavior in healthcare insurance markets[37].

[37] As stated previously, according to the UAE Labor Law, the weekend in the UAE runs from Friday to Saturday for most public services and the government. However, this is not the case in the private sector, for which the weekend runs from Saturday to Sunday. Some workers in the private sector have one day off and work on Sundays, which may suggest the presence of a Sunday effect instead of the Sunday effect. To rule out the potential Sunday effect, we repeat the analysis in which the interaction term between Muslim and Sundays binary variables is the key variable (Sunday*Muslim). The results show that the Sunday effect is positive and distinct among all claim behavior variables. However, when we explore the interaction term between Muslim and Sunday binary variables, the results are insignificant across all claim behavior variables, suggesting that the Sunday effect is not greater among Muslim policyholders than among non-Muslims. Further, we repeat the primary analysis to control for the Saturday effect (as a key variable) in the case where some workers have one day off (usually Fridays) and have to work on Saturday in the private sector. Interestingly, the results show that the Saturday effect is positive and distinct among all claim behavior variables. However, the interaction term between Muslim and Saturday binary variables is negative and significant across all claim behavior variables, suggesting that the Saturday effect is present and that the reduction in asymmetric information is high for Muslims on Saturdays. The results are available upon request.

Table 2.6. Religion and Claim Performance

The table shows the coefficients of interest for the OLS regression results estimating the relationship between our measures of claims behavior and the logged maximum insurance coverage. The key variables of interest are as follows: (1) Muslim, a binary variable equal to one if the policyholder is reported as Muslim and equals 0 otherwise; (2) Muslim*Religious Holiday the (interaction between Muslims and religious holidays); and (3) Muslim*College (the interaction between Muslims and college degree). We include age, gender marital status, relationship to the principal, policy length, chronic condition, and college degree as controls in all regressions. For the control variable descriptions, refer to Table 2-4. The standard errors are reported in parentheses, and *, **, and *** indicate significance at the 10%, 5% and 0.1% levels, respectively. The number of daily observations is provided in square brackets.

	(1)	(2)	(3)	(4)
	Claim Filed	Claim Paid	Amount Filed	Amount Paid
Logged Coverage	0.019***	0.019***	12.596***	11.335***
	(0.000)	(0.000)	(0.361)	(0.332)
Muslim	-0.004***	-0.003***	-2.851***	-2.538***
	(0.000)	(0.000)	(0.157)	(0.143)
Muslim*Religious Holiday	-0.002***	-0.002***	-0.209	-0.182
	(0.000)	(0.000)	(0.205)	(0.188)
Muslim*College	0.001***	0.001***	0.692***	0.710***
	(0.000)	(0.000)	(0.194)	(0.177)
College Degree	0.004***	0.004***	-0.030	-0.103
	(0.000)	(0.000)	(0.168)	(0.154)
Age	0.000***	0.000***	0.113***	0.110***
	(0.000)	(0.000)	(0.005)	(0.005)
Male	-0.002***	-0.002***	-1.613***	-1.376***
	(0.000)	(0.000)	(0.139)	(0.127)
Married	0.004***	0.004***	1.993***	1.834***
	(0.000)	(0.000)	(0.124)	(0.113)
Policy Length	-0.000***	-0.000***	-0.007***	-0.006***
	(0.000)	(0.000)	(0.000)	(0.000)
Spouse	0.011***	0.011***	8.366***	7.609***
	(0.000)	(0.000)	(0.205)	(0.188)
Child Dependent	0.008***	0.008***	4.279***	4.145***
	(0.000)	(0.000)	(0.209)	(0.192)
Chronic	0.022***	0.022***	10.544***	9.568***
	(0.000)	(0.000)	(0.108)	(0.099)
Constant	-0.121***	-0.118***	-104.7***	-94.32***
	(0.00122)	(0.00121)	(2.710)	(2.482)

Observations	33,456,791	33,456,791	33,456,791	33,456,791
R-squared	0.011	0.011	0.001	0.001

Table 2.7: Culture and Claim Performance

The table shows the coefficients of interest for the OLS regression results estimating the relationship between our measures of culture and the logged maximum insurance coverage. The key variables of interest are listed in the rows, and the dependent variables are listed in the columns. For all other culture variable definitions, refer to Table 2-4. We include age, gender, marital status, relationship to the principal, policy length, chronic condition, and college degree as controls in all regressions. The standard errors are reported in parentheses, and *, **, and *** indicate significance at the 10%, 5% and 0.1% levels, respectively. The number of daily observations is provided in square brackets.

	(1)	(2)	(3)	(4)	(5)
	Claim Filed	Claim Paid	Amount Filed	Amount Paid	Logged Coverage
PDI	-0.000***	-0.000***	-0.045***	-0.041***	0.000**
	(0.000)	(0.000)	(0.009)	(0.008)	(0.000)
IDV	-0.000***	-0.000***	-0.033***	-0.031***	0.001***
	(0.000)	(0.000)	(0.006)	(0.006)	(0.000)
MAS	-0.000***	-0.000***	-0.043***	-0.041***	0.001***
	(0.000)	(0.000)	(0.008)	(0.008)	(0.000)
UAI	0.000***	0.000***	0.097***	0.089***	0.001***
	(0.000)	(0.000)	(0.006)	(0.005)	(0.000)
Constant	-0.121***	-0.118***	-104.7***	-94.32***	
	(0.00122)	(0.00121)	(2.710)	(2.482)	
Observations	[29,265,056]	[29,265,056]	[29,265,056]	[29,265,056]	[29,265,056]
R-squared	0.011	0.011	0.001	0.001	

Table 2.8. Friday/Sunday Effects and Claim Performance

The table shows the coefficients of interest in repeating our "primary" analysis (OLS regression results) to estimate the effect of Friday and Sunday on our measures of claiming behavior variables. The key additional variables of interest are as follows: (1) Friday, a binary variable equal to one if the policyholder visits a medical facility on a Friday (Friday effect) and 0 otherwise; (2) Muslim*Friday (the interaction between Muslims and the "religious" day, Friday); (3) Sunday, a binary variable equal to one if the policyholder visits a medical facility on a Sunday (Sunday effect) and that equal 0 otherwise; and (4) Muslim*Sunday (the interaction between Muslims and Sunday). We include age, gender marital status, relationship to the principal, policy length, chronic condition, and college degree as controls in all regressions. For descriptions of the control variables, refer to Table 2.4. The standard errors are reported in parentheses, and *, **, and *** indicate significance at the 10%, 5% and 0.1% levels, respectively. The number of daily observations is provided in square brackets.

	(1)	(2)	(3)	(4)
	Claim Filed	Claim Paid	Amount Filed	Amount Paid
Logged Coverage	0.0106***	0.0103***	8.586***	7.706***
	(9.62e-05)	(9.53e-05)	(0.214)	(0.196)
Muslim	-0.00192***	-0.00182***	-2.025***	-1.834***
	(5.39e-05)	(5.33e-05)	(0.120)	(0.110)
Friday	-0.00968***	-0.00945***	-5.173***	-4.774***
	(9.07e-05)	(8.98e-05)	(0.202)	(0.185)
Friday*Muslim	0.00155***	0.00150***	1.881***	1.797***
	(0.000116)	(0.000115)	(0.259)	(0.237)
Sunday	0.00191***	0.00186***	1.542***	1.307***
	(9.09e-05)	(9.00e-05)	(0.202)	(0.185)
Sunday*Muslim	-0.000463***	-0.000461***	-0.721***	-0.550**
	(0.000117)	(0.000115)	(0.259)	(0.238)
Constant	-0.121***	-0.118***	-104.7***	-94.32***
	(0.00122)	(0.00121)	(2.710)	(2.482)
Observations	[33,456,791]	[33,456,791]	[33,456,791]	[33,456,791]
R-squared	0.011	0.011	0.001	0.001

1.7 Conclusion

Using individual-level insurance data from a Takaful (Islamic) health insurance company located in the UAE, we examine how culture and religion influence asymmetric information. The sample we utilize is unique since individuals in our sample reside in the UAE but originate from 141 different countries. Furthermore, we utilize the reported religious preference and nationality of policyholders to consider how those factors may relate to insurance claiming behavior and the choice of insurance coverage.

Consistent with the hypotheses developed in our theoretical model (see Section 2.3) and economic theory, we find that there is significant evidence of asymmetric information in the Takaful health insurance market. We also find evidence that religious individuals are less likely to file claims and file for claims at smaller amounts than nonreligious individuals. Our results are also evident and consistent with our theoretical model when we control for major Muslim religious holidays (e.g., Eid Al-Fitr, Ramadan, and Eid Al-Adha) and the Sunday effect in insurance markets. The results show that asymmetric information and excessive claiming behavior are significantly lower during the Muslim holidays and Sundays (after the weekend), suggesting that the Sunday effect is present and that the reduction in asymmetric information is pronounced among Muslims, even on Muslim religious holidays and Sundays.

Finally, we find some evidence consistent with our expectation that cultural background is associated with claims behavior and choice in insurance coverage; however, many of these results appear to be economically insignificant. Given the initial results found in this paper that religious individuals engage in less individually advantageous behavior than nonreligious individuals, especially during religious holidays, we believe that more research should be conducted to further deepen our understanding of how culture and religion influence behavior in insurance markets.

References:

Al-Amri, K. (2013). Essays on organizational form and efficiency in the Takaful insurance industry. Temple University.

Ayoub, M. (1991). Islam and Christianity between tolerance and acceptance. *Islam and Christian-Muslim Relations*, *2*(2), 171-181.

Bainbridge, W. S. (1989). The religious ecology of deviance. *American Sociological Review*, 288-295.

Bachman, J., Johnston, L., O'Malley, P. (1993). Monitoring the Future: A Continuing Study of the Life styles and Values of Youth. *University of Michigan Survey Research Center, Ann Arbor*.

Becker, G. S. (1968). Crime and punishment: An economic approach. In The economic dimensions of crime (pp. 13-68). *Palgrave Macmillan, London*.

Browne, M. J., & Doerpinghaus, H. I. (1993). Information asymmetries and adverse selection in the market for individual medical expense insurance. *Journal of Risk and Insurance*, 300-312.

Born, P. H., & Sirmans, E. T. (2019). Restrictive Rating and Adverse Selection in Health Insurance. *Journal of Risk and Insurance*.

Butler, R. J., Gardner, H. H., & Kleinman, N. L. (2013). Workers' compensation: occupational injury insurance's influence on the workplace. In *Handbook of insurance* (pp. 449-469). Springer, New York, NY.

Card, D., & McCall, B. P. (1996). Is workers' compensation covering uninsured medical costs? Evidence from the "Monday effect". *ILR Review*, *49*(4), 690-706.

Campolieti, M., & Hyatt, D. E. (2006). Further evidence on the "Monday Effect" in workers' compensation. *ILR Review*, *59*(3), 438-450.

Campante, F., & Yanagizawa-Drott, D. (2015). Does religion affect economic growth and happiness? Evidence from Ramadan. *The Quarterly Journal of Economics*, *130*(2), 615-658.

Chiappori, P. A., & Salanie, B. (2000). Testing for asymmetric information in insurance markets. *Journal of political Economy*, *108*(1), 56-78.

Chui, A. C., & Kwok, C. C. (2008). National culture and life insurance consumption. *Journal of International Business Studies*, *39*(1), 88-101.

Cohen, A., & Siegelman, P. (2010). Testing for adverse selection in insurance markets. *Journal of Risk and insurance*, *77*(1), 39-84.

Cohen, A., & Einav, L. (2007). Estimating risk preferences from deductible choice. *American economic review*, *97*(3), 745-788.

Cohen, A. (2005). Asymmetric information and learning: Evidence from the automobile insurance market. *Review of Economics and statistics*, *87*(2), 197-207.

Cochran, J. K., & Akers, R. L. (1989). Beyond hellfire: An exploration of the variable effects of religiosity on adolescent marijuana and alcohol use. *Journal of Research in Crime and Delinquency*, *26*(3), 198-225.

Crawford, G. S., Pavanini, N., & Schivardi, F. (2018). Asymmetric information and imperfect competition in lending markets. *American Economic Review*, *108*(7), 1659-1701.

Coolen-Maturi, T. (2013). Islamic insurance (takaful): demand and supply in the UK. *International Journal of Islamic and Middle Eastern Finance and Management*.

Cutler, D. M., & Zeckhauser, R. J. (2000). The anatomy of health insurance. In *Handbook of health economics* (Vol. 1, pp. 563-643). Elsevier.

Cummins, J. D., Smith, B. D., Vance, R. N., & VanDerhei, J. L. (1983). Introduction: Overview of Risk Classification. In *Risk Classification in Life Insurance* (pp. 101-120). Springer, Dordrecht.

Diaz, J. D. (2000). Religion and gambling in sin-city: A statistical analysis of the relationship between religion and gambling patterns in Las Vegas residents. *The Social Science Journal*, *37*(3), 453-458.

Diamond, P. (1992). Organizing the health insurance market. *Econometrica: Journal of the Econometric Society*, 1233-1254.

Doiron, D., Jones, G., & Savage, E. (2008). Healthy, wealthy and insured? The role of self-assessed health in the demand for private health insurance. *Health economics*, *17*(3), 317-334.

Diaz, J. D. (2000). Religion and gambling in sin-city: A statistical analysis of the relationship between religion and gambling patterns in Las Vegas residents. *The Social Science Journal*, *37*(3), 453-458.

Dumitriu, R., & Stefanescu, R. (2017). The Behavior of Stock Prices during Lent and Advent. *Available at SSRN 3092795*.

Eling, M., Jia, R., & Yao, Y. (2017). Between-Group Adverse Selection: Evidence from Group Critical Illness Insurance. *Journal of Risk and Insurance*, *84*(2), 771-809.

Evans, T. D., Cullen, F. T., Dunaway, R. G., & Burton Jr, V. S. (1995). Religion and crime reexamined: The impact of religion, secular controls, and social ecology on adult criminality. *Criminology*, *33*(2), 195-224.

Farley, P. J., & Monheit, A. C. (1985). Selectivity in the demand for health insurance and health care. *Advances in Health Economics and Health Services Research*, *6*, 231-252.

Finkelstein, A., & McGarry, K. (2006). Multiple dimensions of private information: evidence from the long-term care insurance market. *American Economic Review*, *96*(4), 938-958.

Finkelstein, A., & Poterba, J. (2004). Adverse selection in insurance markets: Policyholder evidence from the UK annuity market. *Journal of Political Economy*, *112*(1), 183-208.

Akerlof, G. A. (1970). The Market for Lemons: Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics*.

Guiso, L., Sapienza, P., & Zingales, L. (2003). People's opium? Religion and economic attitudes. *Journal of Monetary Economics*, *50*(1), 225-282.

Guiso, L., Sapienza, P., & Zingales, L. (2009). Cultural biases in economic exchange?. *The Quarterly Journal of Economics*, *124*(3), 1095-1131.

Heaton, T. B., & Pratt, E. L. (1990). The effects of religious homogamy on marital satisfaction and stability. *Journal of Family Issues*, *11*(2), 191-207.

Hilary, G., & Hui, K. W. (2009). Does religion matter in corporate decision making in America?. *Journal of Financial Economics*, *93*(3), 455-473.

Hofstede, G. (1984). *Culture's consequences: International differences in work-related values* (Vol. 5). Sage.

Hussain, M. M., & Pasha, A. T. (2011). Conceptual and operational differences between general takaful and conventional insurance. *Australian Journal of Business and Management Research*, *1*(8), 23-28.

Iyer, S. (2016). The new economics of religion. *Journal of Economic Literature*, *54*(2), 395-441.

Iannaccone, L. R. (1998). Introduction to the Economics of Religion. *Journal of Economic Literature*, *36*(3), 1465-1495.

Kwon, W. J. (2007). Islamic Principle and Takaful Insurance: Re-evaluation. *Journal of Insurance Regulation*, *26*(1).

Lehrer, E. L., & Chiswick, C. U. (1993). Religion as a determinant of marital stability. *Demography*, *30*(3), 385-404.

Long, S. H., Settle, R. F., & Wrightson Jr, C. W. (1988). Employee premiums, availability of alternative plans, and HMO disenrollment. *Medical Care*, 927-938.

Liu, X. (2016). Corruption culture and corporate misconduct. *Journal of Financial Economics*, *122*(2), 307-327.

Maysami, R. C., & Kwon, W. J. (1999). An analysis of Islamic Takaful insurance: A cooperative insurance mechanism. *Journal of Insurance Regulation* 18(1), 109.

Miller, A. S., & Hoffmann, J. P. (1995). Risk and religion: An explanation of gender differences in religiosity. *Journal for the scientific study of religion*, 63-75.

Noussair, C. N., Trautmann, S. T., Van de Kuilen, G., & Vellekoop, N. (2013). Risk aversion and religion. *Journal of Risk and Uncertainty*, *47*(2), 165-183.

Outreville, J. F. (2018). Culture and Life Insurance Ownership: Is It an Issue? *Journal of Insurance Issues*, *41*(2), 168-192.

Park, S. C., & Lemaire, J. (2012). The impact of culture on the demand for non-life insurance. *ASTIN Bulletin: The Journal of the IAA*, *42*(2), 501-527.

Panhans, M. (2019). Adverse selection in ACA exchange markets: evidence from Colorado. *American Economic Journal: Applied Economics*, *11*(2), 1-36.

Pantzalis, C., & Ucar, E. (2014). Religious holidays, investor distraction, and earnings announcement effects. *Journal of Banking & Finance*, *47*, 102-117.

Perrin, R. D. (2000). Religiosity and honesty: Continuing the search for the consequential dimension. *Review of Religious Research*, 534-544.

Phelps, C. (1973). Demand for health insurance: a theoretical and empirical investigation.

Puelz, R., & Snow, A. (1994). Evidence on adverse selection: Equilibrium signaling and cross-subsidization in the insurance market. *Journal of Political Economy*, *102*(2), 236-257.

Rahman, Z. A., & Daud, N. M. (2010). Adverse selection and its consequences on medical and health insurance and takaful in Malaysia. *Humanomics*.

Rothshild, M., & Stiglitz, J. (1976). Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. *quarterly Journal of Economics, 90*(4), 629-647.

Ruser, J. W. (1998). Does workers' compensation encourage hard to diagnose injuries?. *Journal of Risk and Insurance*, 101-124.

Schlesinger, H. (2013). The theory of insurance demand. In *Handbook of insurance* (pp. 167-184). Springer, New York, NY.

Simon, K. I. (2005). Adverse selection in health insurance markets? Evidence from state small-group health insurance reforms. *Journal of Public Economics*, *89*(9-10), 1865-1877.

Smith, A. (1776). The Causes of the Wealth of Nations. *London: W. Strahan and T. Cadell*.

Stavrova, O., & Siegers, P. (2014). Religious prosociality and morality across cultures: How social enforcement of religion shapes the effects of personal religiosity on prosocial and moral attitudes and behaviors. *Personality and Social Psychology Bulletin*, *40*(3), 315-333.

Trinh, T., Nguyen, X., & Sgro, P. (2016). Determinants of non-life insurance expenditure in developed and developing countries: an empirical investigation. *Applied Economics*, *48*(58), 5639-5653.

Wahab, A. R. A., Lewis, M. K., & Hassan, M. K. (2007). Islamic takaful: Business models, Shariah concerns, and proposed solutions. *Thunderbird International Business Review*, *49*(3), 371-396.

Wasiuzzaman, S. (2017). Religious anomalies in Islamic stock markets: the Hajj effect in Saudi Arabia. *Journal of Asset Management*, *18*(3), 157-162.

Wasiuzzaman, S. (2018). Seasonality in the Saudi stock market: The Hajj effect. *The Quarterly Review of Economics and Finance*, *67*, 273-281.

Weber, M. (1905). The Protestant Ethicand the Spirit of Capitalism. *New York: ScribnersWeberThe Protestant Ethic and the Spirit of Capitalism1958*.

Wrightson, J. W., Genuardi, J., & Stephens, S. (1987). Demographic and utilization characteristics of HMO disenrollees. *GHAA journal*, *8*(1), 23-42.

Yaari, M. E. (1965). Uncertain lifetime, life insurance, and the theory of the consumer. *The Review of Economic Studies*, 32(2), 137-150.

Yaari, M. E. (1964). On the consumer's lifetime allocation process. *International Economic Review*, *5*(3), 304-317.

Yousaf, I., Ali, S., & Shah, S. Z. A. (2018). Herding behavior in Ramadan and financial crises: the case of the Pakistani stock market. *Financial Innovation*, *4*(1), 16.

Appendix A. Global Takaful Market Size and Differences between Takaful Insurance and Commercial Insurance

Table A1. Price Differential between Takaful Insurance and Commercial Insurance from 2017 to 2019

Price (USD/Year)	2017	2018	2019
Takaful Insurance	405	412	422
Commercial Insurance	300	305	310

Source: Secondary Sources, Expert Interviews and QYResearch, July 2019

Table A2. Price Differential between Takaful Healthcare Insurance & Commercial Healthcare Insurance Price from 2017 to 2019

Price (USD/Year)		2017	2018	2019
Takaful Healthcare Insurance	Group	600	612	623
	Individual	1200	1232	1240
Commercial Healthcare Insurance	Group	450	455	462
	Individual	901	920	934

Source: Secondary Sources, Expert Interviews and QYResearch, July 2019

Table A3. Global Takaful Market Size by Region 2014-2025 (Million USD)[38]

By Regions	Market Size (Million US$)		CAGR
	2019	2025	(2019-2025)
GCC	20697.7	41687.0	12.38%
Southeast Asia	2823.9	3932.5	5.67%
Africa	1339.1	2786.8	12.99%
Europe	61.7	140.0	14.63%
USA	10.3	21.3	12.87%
China	3.4	5.9	9.62%
Central & South America	2.7	4.5	8.89%
Other	606.9	909.3	6.97%
Total	25545.7	49487.3	11.65%

Source: Secondary Sources, Expert Interviews and QYResearch, July 2019

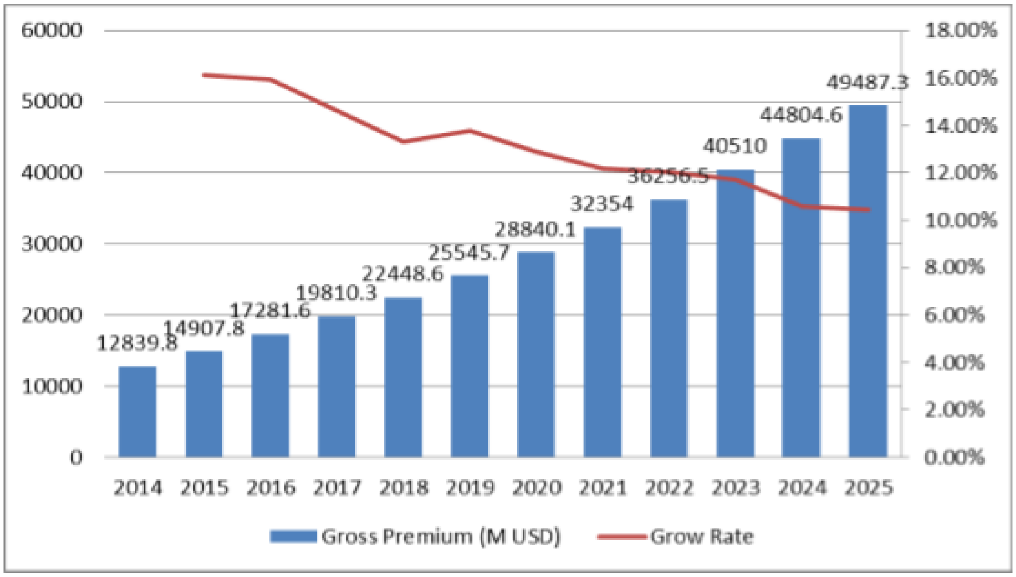
Table A4. Key Difference between Takaful Healthcare Insurance and Commercial Healthcare Insurance

	TAKAFUL	CONVENTIONAL INSURANCE
Definition	A co-operative policy where funds are contributed by donations from participants. The pooled funds can be used to protect other participants from risk.	A policy that shifts the risk to the insurance company. You pay a premium to receive coverage.
Insured Object	Anybody	Anybody
Funds Invested	Operators will only invest in syariah-compliant instruments which are free of gambling, usury and uncertainty	Insurance companies are free to invest in legal instruments like stock, bonds, etc.
Surplus Distribution	It is shared among the participants and operators of a takaful fund	Dividends are returned to shareholders

Source: Secondary Sources, Expert Interviews and QYResearch, July 2019

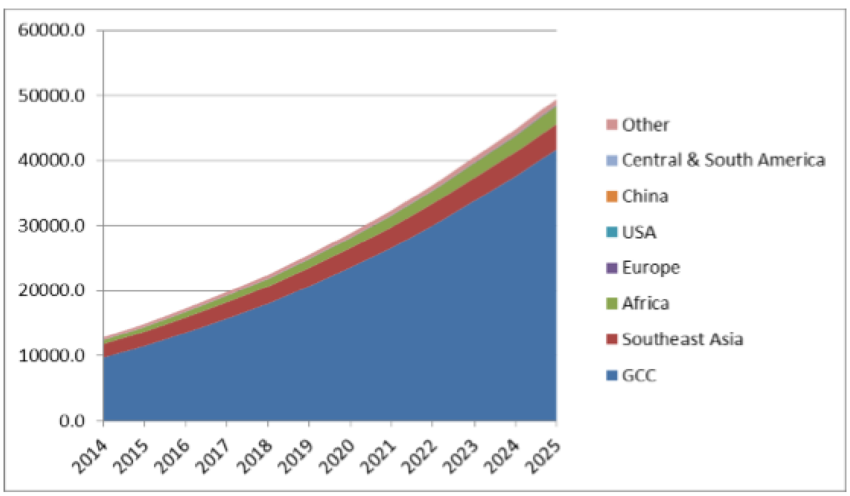
[38] GACR is the compound annual growth rate, and GCC is the Cooperation Council for the Arab States of the Gulf (Saudi Arabia, UAE, Kuwait, Qatar, Bahrain, and Oman).

Figure 2.1. Global Takaful Market Size and Growth Rate 2014-2025 (Million USD)



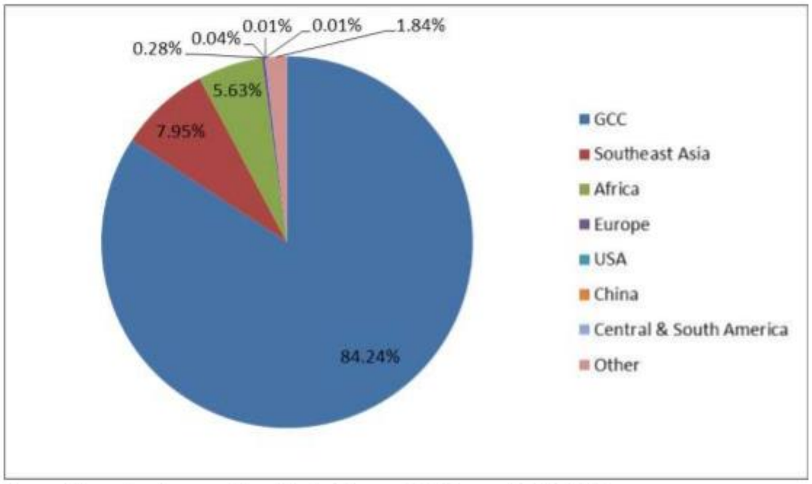
Source: Secondary Sources, Expert Interviews and QYResearch, July 2019

Figure 2.2. Global Takaful Market Size by Region 2014-2025 (Million USD)



Source: Secondary Sources, Expert Interviews and QYResearch, July 2019

Figure 2.3. Global Takaful Market Share by Region 2025



Source: Secondary Sources, Expert Interviews and QYResearch, July 2019

Appendix B. Observations by Country and Assigned Religion for 28 Countries[39]

A2: Observations by Country and Assigned Religion			
Nationality	Individuals	Percent	Muslim Country
India	39,580	42.5	N
Pakistan	11,133	11.95	Y
Bangladesh	6,480	6.96	Y
Philippines	5,409	5.81	N
Egypt	4,122	4.43	Y
United Arab Emirates	3,647	3.92	Y
Jordan	3,198	3.43	Y
Nepal	2,788	2.99	N
Syria	1,924	2.07	Y
Palestine	1,164	1.25	Y
Thailand	1,023	1.1	N
Sri Lanka	752	0.81	N
Lebanon	734	0.79	Y
United Kingdom	615	0.66	N
Sudan	537	0.58	Y
Yemen	433	0.46	Y
United States of America	396	0.43	N
Indonesia	369	0.4	Y
Canada	354	0.38	N
South Africa	313	0.34	N
Iraq	294	0.32	Y
Ethiopia	277	0.3	N
Vietnam	256	0.27	N
South Korea	152	0.16	N
France	147	0.16	N
Australia	136	0.15	N
Morocco	128	0.14	Y
Ukraine	120	0.13	N
Somalia	92	0.1	Y

[39] A list of all other countries is available upon request.

3. Chapter Three: Who Is Responsible? A Study of Potential Fraud in Private Healthcare Insurance Markets

"Workers in the health-care sector need to have high levels of integrity to keep fraud to a minimum."

Lou Saccoccio, CEO at National Health Care Anti-Fraud Association

3.1 Introduction

During recent decades, insurance fraud has been evaluated in many lines of commercial insurance, especially the automobile insurance market (e.g., Tennyson and Salas-Forn 2002, and Derrig 2002). However, Derrig (2002) reports that empirical research on fraud is still limited due to a lack of data[40], especially in healthcare insurance markets. This limitation has naturally led to Ai et al.'s (2018) call to direct future research[41] on healthcare insurance fraud towards private healthcare insurance markets, as insurance economists have been focusing on public and social insurance markets (e.g., Fang and Gong 2017, and Pande and Maas 2013). In this paper, we answer this call by identifying and detecting potential fraud schemes in private healthcare insurance markets; we then examine whether and to what extent fraud is likely to be committed by healthcare providers (HCPs) versus policyholders (PHs).

Fraud in the healthcare insurance market is a pressing concern, as fraudulent healthcare activities are costly. For example, Ai et al. (2018) report that the cost of healthcare insurance fraud amounts to hundreds of billions of dollars each year. Specifically, in 2017, US healthcare spending represented approximately one-sixth (or approximately \$3.5 trillion; 18 percent of the GDP) of the US economy.[42] Thus, it is important to mitigate fraud, waste, and abuse to enhance the efficiency of the healthcare system. Alternatively, as Tennyson (2008) reports, the impact of fraud will affect the overall inequities and inefficiencies in insurance markets. Specifically, inequities occur when the cost of fraud is transformed into a higher premium paid by current PHs. Inefficiencies arise when profiting through defrauding insurers affects the PH's decision to buy insurance and their consciousness of preventing losses (e.g., moral hazard).[43]

---

[40] Most papers on automobile insurance fraud utilized a small dataset that contains approximately 1350 claims (e.g., Artís, Ayuso, and Guillen 1999 and 2002).

[41] The *Journal of Risk and Insurance* published a special issue of five articles about fraud detection in the insurance market in 2002 (Volume 69, Issue 3) and called for future research in this area.

[42] See the website of the Committee for Responsible Federal Budget (https://www.crfb.org/papers/american-health-care-health-spending-and-federal-budget)

[43] Artís, Ayuso, and Guillen (1999) state that fraud is a prime example of moral hazard and that controlling fraudulent schemes requires a deep knowledge and understanding of the PHs' behavior.

Interestingly, healthcare fraud seems to be much more complicated and unique than automobile insurance fraud for two major reasons. (1) Health insurance contracts are unique in that PHs are generally unable to submit a claim unless the action is accompanied by a submission from an HCP. Health insurance is in contrast to other lines of insurance, such as homeowners or automobile insurance, in which PHs generally file the claim directly with the insurer. (2) Most importantly, unlike automobile insurance fraud, healthcare fraud can be initiated by any of multiple parties, including the beneficiary, the provider, and/or any other intermediaries. Johnson and Nagarur (2015) shed light on potential fraud schemes by physicians and patients. They clarify that HCPs (i.e., doctors) may provide unnecessary services or prescribe irrelevant medications to increase their revenues. Patients may also commit fraud by filing claims for services that never happened or claiming under someone else's insurance policy. These results suggest that opportunities for committing fraud in the healthcare insurance market are very different from those in other insurance markets. Therefore, the cost of handling medical claims may be significantly higher. Consequently, it is worth identifying and detecting fraud in healthcare insurance markets, as the evidence may shed light on the best avenues for addressing HCP versus PH fraud schemes and the cost of handling medical claims.

Given the nature of healthcare insurance markets and health insurance contracts, we argue that compared to PHs, HCPs are presented with more opportunities to behave dishonestly and defraud insurance companies. Consistent with our expectation, we find the following four results. First, fraud/abuse in the healthcare insurance market is committed largely by HCPs rather than by PHs. Specifically, we show that fraud/abuse from the PH side is approximately 40 percent lower than that from the HCP side.[44] Second, fraudulent claims are more likely to be rejected when the HCP rather than the PH controls the claim. Third, using principal component analysis of RIDIT scores (PRIDIT), we find that the fraud "suspicion" rate is lower in the inpatient claims group than in the outpatient claims group. Fourth, we show that there is no significant difference between PRIDIT scores (as a proxy for suspiciousness) when we consider insurance type (i.e., Takaful and commercial insurance). Our PRIDIT results reveal that there is no significant difference in fraud patterns between the Takaful (Islamic) and commercial healthcare insurance markets. Finally, we show that there is no significant difference between suspicious claims when we consider the policyholders' religious affiliation, and culture. Given that the Islamic and commercial healthcare insurance markets are similar in that HCPs (versus PHs) have more control over

[44] The underlying economic rationale is that HCPs misuse their patients' insurance in order to obtain money, especially when the probability of being caught is low.

claims, our fourth result further resonates with our main finding that fraud in the healthcare insurance market is committed largely by HCPs.

This paper makes the following two contributions. First, to the best of our knowledge, it is the first study to identify and detect fraud schemes in two private health insurance markets. Specifically, it is the first to provide empirical support that healthcare fraud is more pronounced among HCPs than among PHs. Second, this paper contributes to the existing literature on insurance fraud by identifying and detecting potential fraud schemes using a large sample of claims (our dataset contains 633,042 claims versus 1,350 claims in Artis, Ayuso and Guillen (2002)).[45] The uniqueness of our data allows us to answer the call of Ai, Lieberthal, Smith, and Wojciechowsk (2018) to direct the focus of the healthcare insurance fraud literature more towards the private insurance markets and the factors and circumstances that are likely to influence insurance fraud.[46]

Thus, given our results, evaluating and understanding the extent and potential drivers of fraud provide public policy implications for all parties in the insurance industry: policymakers (insurers), PHs, and regulators. For example, regulators can set and/or update current regulations to better control the PHs' and the HCPs' fraudulent actions. Policymakers can develop better strategies to investigate claims fairly and settle them efficiently. Furthermore, the results of our study and other healthcare fraud studies will help policymakers quantify the healthcare fraud rate and charge more accurate premiums (Ai et al., 2018). As a result, PHs will benefit from a more efficient healthcare insurance market by being able to pay a fairer premium.[47]

The remainder of this paper will proceed as follows. The second section provides a literature review for our analysis, followed by a discussion of our hypothesis. The third section provides a detailed description of the data and methodology used to test our hypotheses (probit model analysis). The fourth section provides a discussion of the probit regression results. The fifth and sixth sections present the PRIDIT model analysis and its results. The seventh section provides an examination of classifying claims by PRIDIT scores, including which claims should be audited. The final section provides concluding remarks and suggestions for future research.

[45] For instance, Artís, Ayuso, and Guillen (2002) conduct research on automobile insurance fraud by utilizing a small dataset that contains approximately 1350 claims.
[46] The authors state, "The prevalence of fraud, and how it differs by service type, country or public versus private insurance, is a key gap in the literature that should be addressed. Validated fraud studies on a small scale can be used to generate evidence on the prevalence of fraud that can be extrapolated or used to power larger-scale studies of health care fraud."
[47] For more information, see Tennyson's (2008) discussion of the impact of fraud on the overall inequities and inefficiencies in insurance markets.

3.2 Literature Review and Hypothesis Development

3.2.1 Introduction and Operational Definition of Healthcare Fraud

Fraud detection is complex, especially in healthcare claims, because healthcare fraud can be committed by various parties, including the PH (patient), the HCP, and any intermediaries. It is essential to understand the definition of fraud, as the definitions fall on a broad spectrum. For example, the definitions range from waste and abuse to fraud, with fraud at the extreme end. In the healthcare insurance literature, Ai et al. (2018) define fraud as "*…an intentional deception or misrepresentation made by a person with the knowledge that the deception could result in some unauthorized benefit to himself or to some other person*."

For all players in the healthcare insurance market, Berwick and Hackbarth (2012) further classify the term "waste" in the following six categories: (1) pricing failures, (2) failures of care coordination, (3) overtreatment, (4) administrative complexity, (5) failures in the execution of care processes, and (6) fraud and abuse. Thus, to reduce waste, abuse, and fraud, it is imperative to understand how players in the healthcare market (especially HCPs, PHs, and insurance carriers) perceive medical services and to determine their place on the continuous scale of waste, abuse, and fraud. For instance, giving medications to a relative after being treated might be legitimate from a patient's perspective. Nevertheless, this action is considered abuse and/or misuse of insurance from an insurer's perspective. In this study, we cover all types of "waste, abuse, and fraud by HCPs and PHs in the healthcare insurance market, with fraud at the extreme end of the spectrum."

3.2.2 Healthcare Fraud Literature

Recent insurance fraud research falls into two main areas: (1) automobile and workers' compensation insurance fraud (e.g., Weisberg and Derrig, 1991; Weisberg and Derrig, 1992 and 1996) and (2) healthcare fraud (e.g., Parente et al. 2011 and Ai et al. 2018). As this study is on healthcare fraud, we will focus on the healthcare fraud literature.[48]

Healthcare insurance fraud is complicated, as it can be initiated by multiple parties, including the beneficiary, the provider, and/or any intermediaries. Furthermore, the nature and structure of healthcare insurance claims are quite different. For example, at each medical appointment, various medical services

[48] For more information about automobile and workers' compensation insurance fraud, see the five insurance fraud articles that were included in the special issue on insurance fraud detection and those that have been published in the *Journal of Risk and Insurance* since 2002.

may be provided and are likely to be filed separately.[49] Due to a lack of data, this situation makes healthcare fraud investigations unique and limited.

This situation naturally motivated Ai et al. (2018) to conduct a systematic literature review of healthcare insurance fraud; their findings reveal that insurance economists have been utilizing data from public and social insurance markets (e.g., Medicare and Medicaid) within the US and other countries. In their analysis, the authors include only twenty-seven papers, of which most were from health services research or from computing and information services journals (e.g., Liou et al. 2008, and Johnson and Nagarur, 2016). Furthermore, the majority of studies were on medical services only, and data clustering was the primary research approach.[50]

According to Ai et al. (2018), only five papers have reported the prevalence of healthcare insurance fraud. These five studies are most closely related to our article, and they reveal fraud rates that support our argument that fraud/abuse in the healthcare insurance market is more likely to be committed by HCPs than by PHs. For instance, Joudaki et al. (2016) examine healthcare fraud within the drug filings of general medical doctors. The authors identify and detect fraud/abuse schemes among doctors who are likely to commit such abusive behaviors. The results reveal that 54 percent of medical doctors are suspected of abusing and misusing insurance and that only 2 percent of medical doctors are suspected of committing fraud. In the same context, healthcare fraud is prevalent in public and social insurance programs. For example, van Capelleveen et al. (2016) examine healthcare fraud in the Medicaid claims of 369 primary dental providers and reveal that 71 percent of the top dental providers (12 of the 17 top providers) filed abuse claims.

Shin et al. (2012) utilize a scoring model to detect fraud patterns among Korean internal medicine clinics. Consistent with Capelleveen et al. (2016), the authors report that fraud/abuse is committed mostly by HCPs and that approximately 6% of internal medicine clinics are suspicious, suggesting abusive utilization behaviors. Finally, Major and Riedinger (2002) utilize electronic fraud detection (EFD) to identify healthcare insurance fraud and report that approximately 4 percent of HCPs are brought to the attention of fraud units because of fraudulent activities.

[49] Farbmacher and Spindler (2019) state that *"…each claim consists of multiple items, including the dates of treatment, a short description of the task or tasks performed, and the associated costs. The number of items varies from patient to patient, resulting in claims of variable length"*

[50] The authors report that one major problem with existing fraud detection methods is that audit-based approaches may not be validated or may rely on particular human experts. The authors conclude that the main gaps in the existing fraud research are (1) a need for validation of the existing methods, (2) a lack of proof of intent to commit fraud, and (3) inadequate estimations of the fraud rate for many programs.

In the healthcare insurance market, two more severe issues that are distinct from but related to healthcare insurance fraud, especially HCP fraud, are *overbilling and upcoding*.[51] For instance, upcoding occurs when HCPs change the codes of provided medical services to more intensive and expensive services to gain more profit from insurance companies. Healthcare insurance economists have conducted research to address this type of fraud in the context of Medicare.[52] For instance, Geruso and Layton (2015) propose a theoretical model to evaluate the effect of upcoding within private health plans versus comparable public health plans. The authors report that because of diagnosis upcoding, the US federal government issues annual overpayments that amount to $50 billion. In the same context, Goates (2010) examines the relationship between upcoding, insurance type, and disease type (international classification of diseases, ICD9) and finds that the probability of being prescribed more drugs is significantly lower when the patient is not enrolled in a Medicare Prospective Payment System (PPS).[53] These results suggest that instead of using the entire sample, insurance economists should focus their fraud analysis on the most heavily utilized drugs or the drugs with more substantial claim amounts.

An additional study on overbilling in the context of Medicare is conducted by Fang and Gong (2016), who detect potential overbilling by utilizing the number of hours physicians worked on providing medical services and submitting service codes to Medicare.[54] They find that approximately 3 percent of Medicare Part B medical doctors (2,300 physicians) file more than 100 hours per week to Medicare for reimbursement.[55] Interestingly, the authors find that the flagged physicians are more responsive to economic gain, filing codes with more intensive medical services than the codes filed by unflagged physicians. They also find that this 3 percent of doctors is more likely to perform such fraudulent/abusive behaviors, especially if the financial gain from filing more intensive medical codes is quite high.

Given the results reported by Ai et al. (2018) and Cremeans et al. (2019), we expect that fraud and abusive behaviors in private healthcare insurance markets are likely to be committed by HCPs. However,

---

[51] Johnson and Nagarur (2015) shed light on potential fraud schemes by physicians and patients. They clarify that HCPs, e.g., doctors, may provide unnecessary medical services, prescribe irrelevant medications, or change the codes of provided medical services to increase revenues.

[52] Cremeans et al. (2019) conduct a systematic literature review of physician upcoding and Medicare fraud reported in research articles published from 2005 through 2018. They review 22 research articles to examine abuse, Medicare fraud, and, most importantly, the effect of upcoding on Medicare expenditures. They state that with the significant increase in upcoding, fraudulent Medicare charges have reached USD $12.5 billion since 2007 and, most importantly, have caused substantial increases in Medicare overpayments.

[53] The authors state that a 10% increase in Medicare patient admissions at medical facilities leads to a 1.1 percent increase in the probability of being prescribed a higher amount of drugs.

[54] The authors define overbilling as a situation characterized by the following: "…providers report more and/or higher intensity service codes than actually delivered to receive higher Medicare reimbursement."

[55] The authors state that the accuracy of these hours is doubtful, as National Ambulatory Medical Care Survey data reveal that HCPs spend a maximum of 50 hours per week treating Medicare patients.

regardless of whether fraud/abuse is committed by HCPs and/or PHs, in such cases, business common sense suggests that such medical claims will be rejected. Fraudulent claims are likely to be dismissed due to fraudulent/abuse actions from the HCP side, the PH side, and/or both (the HCP/PH side) when the HCP and PH collude. Across these three categories, the level of fraud/abuse and the motivation to commit it might differ depending on each individual's honesty and integrity.[56]

Picard (1996) assumes that insurance buyers are either honest or opportunistic, and Table 3.1 shows the PH and HCP types and the potential scenarios of collusion between them. We expect the lowest fraud rate to occur when both players are honest, as the probability of collusion between them is small. However, if two opportunistic players collude with one another, we expect the highest fraud rate. *Our theoretical tension is represented by the scenario in which an opportunistic HCP treats an honest PH.[57] While honest PHs can mitigate healthcare fraud by not colluding with HCPs, the fraud rate is expected to be high, as there is no control over the HCP's abusive behaviors (e.g., the HCP will file claims for unnecessary medical treatments).*

Table 3.1. Policyholder and Healthcare Provider Types and Potential Collusion Scenarios

Policyholders		Healthcare Providers	
		Honest (H)	***Opportunistic (O)***

[56] The economic rationale behind healthcare fraud suggests that certain HCPs are likely to misrepresent the provided medical services in order to profit if the probability of detection is low. For more information, see Ai et al. (2018) and Cremeans et al. (2019).

[57] We must note that our data do not contain information about HCP and PH integrity and whether they are honest and/or o*pportunistic*. However, this discussion (Table 1) helps us to understand the extent of healthcare insurance fraud, especially abusive/fraudulent behaviors from the HCP side. For more information, see the data subsection (1.3.1).

	Honest (H)	$H_{PH}H_{HCP}$: Lowest fraud rate; close to zero (No collusion)	$H_{PH}O_{HCP:}$ Fraud rate from the HCP side is high. However, honest policyholders can mitigate it (No collusion)
	Opportunistic (O)	$O_{PH}H_{HCP:}$ Fraud rate from PH side is low; close to zero. Honest HCPs can significantly mitigate it (No collusion)	$O_{PH}O_{HCP}$: Highest fraud rate. Significant probability of collusion between HCP and PH
	Summary: $H_{PH}H_{HCP} < O_{PH}H_{HP} < H_{PH}O_{HCP} < O_{PH}O_{HCP}$		

We previously noted (see Section 1.3.3 in the first essay) that either through risk misrepresentation (underwriting) or insurance fraud, risk-averse PHs are presented with an opportunity to behave dishonestly. The same argument applies to HCPs. However, given the nature of healthcare insurance, the difference between the two situations is that HCPs are presented with more opportunities to behave dishonestly and defraud insurance companies, as by not only providing medical services but also filing claims directly with insurers, the HCPs are more involved with the medical process than the PHs, . This aspect is likely to mitigate fraud from the PH side, as one would expect. However, there is no control over abusive/fraudulent behaviors on the HCP side. Therefore, we formally hypothesize the following.

H1. Fraud/abuse in the healthcare insurance market is committed mainly by HCPs, and fraudulent claims are more likely to be rejected when the HCPs have more control of claims than the PHs.

Parente et al. (2012) utilize Medicare claims to identify healthcare fraud and extrapolate the prevalence of fraud and abuse in the healthcare insurance market. They find a lower rate of fraud in inpatient medical claims than in outpatient medical claims. This makes sense, as outpatient medical visits are likely to be suspicious because PHs might be well connected with medical doctors outside the medical network and visit those doctors only to collude with them and defraud insurance companies for financial gain.[58] Given this result, we formally postulate the following.

[58] We acknowledge that some policyholders might be honest and seeking specialized medical doctors for better medical treatment, these policyholders are likely to visit a medical doctor outside the medical network if they suffer from critical diseases.

H2. The fraud "suspicion" rate is lower in the inpatient claims group than in the outpatient claims group, particularly in the case of claims from HCPs outside the medical network.

3.3 Data and Methodology: Probit Model Analysis

3.3.1 Data for Probit Analysis

To test our hypotheses, we use a proprietary dataset that contains all PHs and individual medical claims from two private insurance companies (Takaful and Commercial) operating in the United Arab Emirates (UAE) and Jordan[59] from 2013 to February 2015.

Our data are useful for studying healthcare fraud for the following two reasons.

- *First*, our dataset contains fraudulent claims reported by the fraud department within the insurance company (10,236 claims). The insurance company also provided us with information about whether fraud/abuse was committed by the HCP (HCP fraud), PH (PH fraud), or both when colluding to gain from insurance (HCP_PH fraud). This dataset enables us to examine whether fraud/abuse in the healthcare insurance market is primarily committed by HCPs and whether fraudulent claims are more likely to be rejected when the HCP rather than the PH controls the claim (H1). *Furthermore, because we can observe the investigator's comments and process codes for all fraudulent rejected claims in the fraudulent claims file, these data enable us to report healthcare fraud schemes.*[60]
- *Second*, this dataset contains the characteristics of each PH's insurance policy (i.e., policy commencement and termination dates), medical visit characteristics (i.e., HCP type, disease type (ICD9 codes) and claimed items), and policyholder characteristics (i.e., age, gender, dependents, and marital status). With these unique variables available to us, we can examine our hypotheses by simultaneously controlling for the effects of coverage type, dependent type, HCP type, disease type (ICD9 category), gender, marital status, and age.

Most of the claims in our dataset were either fully covered with a process code (A00) or adjusted to the coinsurance and/or agreed price with a process code (C01). The remainder were rejected as fraudulent (10,237 claims), and 28 different processing codes were used to indicate the fraudster side (e.g., HCP vs. PH fraud), fraud *schemes*, and the rejection reasons. For the investigation of the first hypothesis, to

---

[59] The two companies are representative of the UAE and Jordanian insurance markets, as they have a large market share among Takaful and commercial companies.

[60] We must acknowledge that Takaful and commercial insurance companies have different business models, which may influence their strategies of auditing claims, identifying fraud, and/or settling medical claims (for more information, see Section 1.2 (the Takaful insurance section) in essay one.

examine whether fraud/abuse in the healthcare insurance market is committed largely by HCPs, we first utilize the fraud report and the investigator's explanation for all fraud *schemes* to create several measures of fraudsters. We then examine whether fraudulent claims are more likely to be rejected when the HCP rather than the PH controls the claim. Artís et al. (1999) use fully rejected claims and the policies deleted by insurers as a basis for identifying fraud. Our approach is superior, as we use the actual insurance company's fraud report, which recognizes that the insurer's classification of fraudulent claims might be biased and that some of the rejected claims may be due to mistakes on the part of the PHs or the HCPs. According to the insurer's fraud report, (16) processing codes and counts (7,797 claims) are connected to HCP fraud, and (7) processing codes and counts (1,690 claims) are connected to PH fraud. Finally, (3) processing codes and counts (749 claims) are connected to HCP/PH fraud. The frequency of this claim disposition and the investigator's classification of fraud *schemes,* along with the processing codes, processing results, and the investigator's explanation for all fraud *schemes,* are shown in Table 3.2.

Table 3.2 reveals that the most common fraud and/or abuse *schemes* from the HCP side are (1) overutilizing services, (2) providing unnecessary services, (3) misrepresenting the dates of service, (4) prescribing unregistered alternative medications to receive referral fees from pharmaceutical companies, and (5) upcoding.[61] The fraud schemes from the PH side include (1) corruption, (2) medical identity theft, and (3) the misrepresentation of dates/locations of service. To a large extent, such fraud schemes are consistent with the findings of Johnson and Nagarur (2015), who report potential fraud schemes by physicians and patients.[62] Most importantly, some fraud is committed due to collusion between the HCPs and the PHs. For instance, false claims are filed to refill prescriptions before the expiration of old dosages. Cummins and Tennyson (1996) state that dishonesty in insurance claims is present in the insurance market, with PHs inflating the size of their claims to compensate for deductibles, copayments, and paid premiums. For example, they note that PHs can defraud insurance companies by "…allowing a doctor to submit medical bills for treatment not received to increase the insurance settlement amount, and being involved with an organized ring of doctors, lawyers, or body shops that file false claims to get money from insurance companies."

---

[61] Upcoding, in which HCPs change the codes of provided services to more intensive and expensive services to profit from insurance companies, is a serious issue in the healthcare insurance market. Healthcare insurance economists have conducted research to address this type of fraud in in the context of Medicare. For more information see Geruso and Layton (2015), and Cremeans et al. (2019).

[62] For example, the authors clarify that HCPs, e.g., doctors, may provide unnecessary services or prescribe irrelevant medications. Patients may also commit fraud by misrepresenting themselves and/or providing irrelevant information to insurers. Patients can use another person's insurance policy, file claims in which the policy does not cover subscriber dependents, and file a claim for a service provided after the expiration of the policy. For more information, see Johnson and Nagarur (2015).

Table 3.2. Claim Disposition and Classification of Fraudulent Claims

Process code	Frequency	Fraudster Side	Process Results	Investigator's Comments
A00	126,806 (20.96%)	No fraud was reported	Claim accepted fully	None
C01	467,920 (77.34%)	No fraud was reported	Claim adjusted according to coinsurance	None
R01	111	No fraud was reported	Claim correction (Claim corrected to be fully covered)	None
O03	3273	HCP	Claim rejected – The beneficiary is not entitled to be covered for more than 12 per policy or per subscriber	HCP abuse
L05	1565	HCP	Claim rejected – The claim should arrive within 45 days of the claim date	HCP fraud: Misrepresenting dates of service
T03/TO3	1075	HCP	Claim rejected – The disease is not covered by the policy	HCP fraud: Billing for a non-covered disease/service as a covered disease/service
T26	602	HCP	Claim rejected – The medicine is not officially registered with the Ministry of Health	HCP fraud: HCP prescribing alternative unregistered medications to receive payment for referrals from pharmaceutical companies
O02	450	HCP	Claim rejected – Item and/or service not covered by the beneficiary's policy	HCP fraud: Overutilization of services
T21	400	HCP	Claim rejected – The claimed procedure is not prescribed by the specialist responsible for the beneficiary	HCP abuse: Misrepresenting the provider of service (reporting of procedures by unspecialized medical doctor)
T14	192	HCP	Claim rejected – The illness or disease procedure cannot occur more than 1 time in the beneficiary's lifetime	HCP fraud: Providing unnecessary services/incorrect reporting of diagnoses and/or procedures
V01	103	HCP	Claim rejected – The follow-up visit is free of charge within 72 hours	HCP fraud: Upcoding the follow-up visit/misrepresenting the type of service

H03	83	HCP	Claim rejected – The doctor or HCP cannot claim for specific medications or services	HCP abuse: Insurance misuse
V00	67	HCP	Claim rejected – The beneficiary cannot refer to the same specialist twice within 72 hours	HCP abuse: Overutilization of services/false or unnecessary medical visit
I03	18	HCP	Claim rejected – The third-party administrator (TPA) did not authorize this visit	HCP fraud: Billing for a nonauthorized service as an authorized service
H02	15	HCP	Claim rejected – The HCP is suspended	HCP abuse
T10	6	HCP	Claim rejected – The medicine cannot affect people under 12 years old.	HCP fraud: Unnecessary issuance of prescription drugs
T13	4	HCP	Claim rejected – No processing results	HCP fraud: Upcoding/unbundling
O00	1	HCP	Claim rejected – The outpatient visit is not covered for first coverage class.	HCP abuse
T00	1	HCP	Claim rejected – Prescribing some kinds of medicine is not allowed by a dentist	HCP abuse: The procedures/medications prescribed are not allowed by a dentist
E12	1566	PH	Claim rejected – The beneficiary is suspended	PH fraud: Corruption/fraud
E09	26	PH	Claim rejected – Married daughters are not covered	PH fraud: Misrepresenting the marital status of married daughters
L03	24	PH	Claim rejected – The date of visit is before the date of subscriber enrollment	PH fraud: Misrepresenting the date of the subscriber visit
E02	21	PH	Claim rejected – The policy does not cover subscriber dependents	PH fraud: Billing for nondependents as covered dependents/medical identity theft
E06	5	PH	Claim rejected – Multiple wives are not covered by the head of family (HOF) policy	PH fraud: Incorrect reporting of the covered wife
V02	5	PH	The policy is not valid in Egypt or Palestine	PH fraud: Misrepresenting location of service
L02	3	PH	Claim rejected – The policy expired before the date of service	PH fraud: Misrepresenting dates of service.
T06	645	HCP/PH	Claim rejected – The disease coverage has a special waiting period, and the disease is not covered by the policy before 365 days have passed or until 280 days after the enrollment date	Collusion of PH/HCP: PH misrepresentation and misuse of insurance by the HCP/PH.

T20	36	HCP/PH	Claim rejected – The beneficiary is not entitled to refill his/her medicine dosage, as his/her old dosage has not yet expired	Collusion of PH/HCP: Overutilization of services from the HCP and the PH sides
T15	1	HCP/PH	Claim rejected – The TOTAL procedure cannot be applied for a female	HCP/PH fraud: Corruption and collusion between the patient and the medical doctor
Total	**605,037**			

The variables utilized in this paper are listed in Table 3-3 along with brief descriptions. We use the following three measures of healthcare fraud: (1) the number of fraudulent claims filed by the HCP; (2) the number of fraudulent claims filed by the PH; and (3) the number of fraudulent claims identified as collusion between HCPs and PHs.

Table 3.3. Description of Variables

Variable	Description
Dependent variables	
Fraudulent_Claims	1 if the claim is fraudulent (reported by the insurer)
Key variables of interest	
HCP_Fraud	1 if fraud is committed from the healthcare provider (HCP) side (the admitted variable)
PH_Fraud	1 if fraud is committed from the policyholder (PH) side
HCP_PHFraud	1 if fraud is committed due to collusion of PH and HCP
Control Variables	
(1) Healthcare Provider Type	
Pharmacy	1 if the HCP type is pharmacy (the admitted variable)
Doctor	1 if the HCP type is a medical doctor
Emergency Center	1 if the HCP type is a medical emergency center
Inpatient Facility	1 if the HCP type is an inpatient facility
Lab	1 if the HCP type is a medical lab
Radiology Center	1 if the HCP type is a radiology center
Medical Service Center	1 if the HCP type is a medical service center
(2) Policyholder Characteristics	
Age	Age of beneficiary
Male	1 if male and 0 otherwise; male is the admitted variable
Married	1 if married; single is the admitted variable
Principal	1 if the principal of the policy; principal is the admitted variable
Spouse	1 if the se spouse of the PH
Child Dependent	1 if the child dependent of the PH
(3) Coverage Types	
Suite Class	1 if the PH coverage type is suite/executive (the most generous medical class)
Third_Class	1 if the PH coverage type is third class (the second most generous medical class)
Second_Class	1 if the PH coverage type is second class (the basic medical class)
First_Class	1 if the PH coverage type is first class (the least generous medical class); first class is the admitted variable
(4) Classification of Diseases (ICD9 Chapters)	
ICD9_Cat1	1 for diseases of the respiratory system: list of ICD-9 codes 460–519; ICD9_Cat1 is the admitted variable
ICD9_Cat2	1 for infectious and parasitic diseases: list of ICD-9 codes 001–139
ICD9_Cat3	1 for dermatologic, musculoskeletal and rheumatologic disorders: list of ICD-9 codes 680–709 and ICD-9 codes 710–739

ICD9_Cat4	1 for symptoms, signs and ill-defined conditions: list of ICD-9 codes 780–799
ICD9_Cat5	1 for injury and poisoning diseases: list of ICD-9 codes 800–99 and list of ICD-9 E and V codes
ICD9_Cat6	1 for neoplasms, hematology and oncology disorders: list of ICD-9 codes 140–239 and ICD-9 codes 280–289
ICD9_Cat7	1 for endocrine, nutritional and metabolic diseases, and immunity disorders: list of ICD-9 codes 240–279
ICD9_Cat8	1 for mental/psychiatric disorders and diseases of the nervous system and sense organs: list of ICD-9 codes 290–319, ICD-9 codes 320–359, and ICD-9 codes 360–389
ICD9_Cat9	1 for diseases of the circulatory system: list of ICD-9 codes 390–459
ICD9_Cat10	1 for diseases of the digestive system: list of ICD-9 codes 520–579
ICD9_Cat11	1 for diseases of the genitourinary system: list of ICD-9 codes 580–629
ICD9_Cat12	1 for complications of pregnancy, childbirth, and the puerperium and 1 for congenital anomalies and certain conditions originating in the perinatal period: list of ICD-9 codes 630–679, ICD-9 codes 740–759, and ICD-9 codes 760–779
(6) Claim Disposition/Characteristics	
Fully_Covered	1 if the claim is fully covered
Adjusted_Claims	1 if claims adjusted to coinsurance and the agreed price (the admitted variable)
Fraudulent_Claims	1 if the claim is fraudulent (reported by the insurer)
Coinsurance	Amount of money adjusted to coinsurance and agreed price
Due_Value	Amount of money paid to the beneficiary
Rejected	Rejected amount of the submitted claim
File_Days	Claim filing period: the difference between the claim's received date and visit date (in days)
Medical_Network	1 if the PH visits the HCP from the medical network and 0 otherwise (within network is the omitted variable)

3.3.2 Methodology

Our first strategy is to use a probit model to test our core Hypothesis 1 regarding whether fraud/abuse in the healthcare insurance market is committed largely by HCPs rather than by PHs and whether fraudulent claims are more likely to be rejected when the HCP rather than the PH controls the claim. Fraud from HCPs is the base and is expected to be larger than fraud from PHs or than fraud when there is collusion between HCPs and PHs. Furthermore, the likelihood of rejecting fraudulent claims when the HCP controls the claim is expected to be larger than in other cases.

Following Cummings and Tennyson (1996) and Bhaskaran and Puelz (2009), we simultaneously consider the following eight control variables: (1) HCP type for every medical visit, (2) gender, (3) age, (4) relationship to the principal PH (i.e., spouse or child), (4) marital status, (5) coverage type for each beneficiary under an insurance policy, (6) diagnosed diseases for every medical visit (ICD9), (7) HCP medical network, and (8) delay in filing claims. Cummins and Tennyson (1996) state that such variables are more likely to affect the filing of both valid and fraudulent claims. Furthermore, Bhaskaran and Puelz (2009) report that in settling claims and/or committing fraud, especially in the settlement size and settlement time, the nature of the claim, PH characteristics, and attorney involvement are crucial control factors.

Given Bhaskaran and Puelz's (2009) results, claim size as a control variable is very important in our analysis. We argue that claim size is significantly associated with claim disposition, HCP type, the type of medical service (ICD9), and coverage type. It is vital to observe the variations in claim size across claim disposition, HCP type, the type of service, and class type. For instance, Figures 3.1 and 3.2 represent the variation in claim size across HCP type and coverage type, respectively. Because some HCP types and certain medical coverage classes provide intensive and better-quality services, it is important to observe whether the distribution of such claims differs. Figure 3.1 reveals that the size of the claim clearly varies among HCP types. The distribution of claims in which medical services were provided by inpatient facilities (4), medical service centers (7), and radiology centers (6) was much larger than that in which medical services were provided by other HCPs. This variation makes sense since such medical facilities provide more intense services. In the same context, Figure 3.2 shows that medical coverage classes that provide generous and intense medical services have much larger claims than others. This is obvious in our data, since the suite class is the most generous medical coverage and has the largest distribution of the coverage classes.

To show the variation in claim size and settlement period across disease types, we utilize the ICD9 codes and ICD9 description to group all diseases into 12 categories.[63] Figure 3.3 represents the distribution of claim size and settlement period across all ICD9 chapters. According to Figure 3.3, claim size varies significantly across all ICD9 chapters, and *C9* (complications in pregnancy, childbirth, and puerperium) has the largest distribution amount among all categories.

A longer settlement period might provide a signal of whether a claim is fraudulent. When insurers receive claims, they spend time reviewing them until they can settle them fairly. Insurance companies verify claims through investigation or auditing to ensure that they are valid claims. During the verification process, insurers utilize all observable claim/PH characteristics (e.g., claim size, provider type, medical

[63] In consultation with a physician specializing in internal medicine/endocrinology and diabetes, we regrouped all similar diseases into 12 categories of services/disease areas.

coverage type) and other factors, including the type of medical service[64]. Insurers spend more or less time, depending on the case. If insurers suspect that certain claims are invalid or suspicious, they are likely to spend more time investigating them. A longer settlement time in such cases enables insurers to obtain more information about the claims and therefore provide a fair offer based on the claimant's actual economic loss. After the verification process, claims are settled as follows: fully covered; fully rejected; adjusted to the coinsurance, deductibles and/or agreed price; or partially denied. This finding suggests that settlement time might be associated with many factors, including claim size, claim disposition, provider type, disease type, dependent type, visit type and class type.

Figure 3.3 also represents the distribution of settlement time across ICD9 categories and shows an apparent variation in the settlement period across all ICD9 chapters. Most importantly, it reveals that the settlement period for PHs diagnosed with diseases of the circulatory system (C5) and endocrine, nutritional and metabolic diseases (C3) is significantly longer than the period for other ICD9 categories. This finding supports our earlier argument that insurers may spend more time investigating more substantial claims and/or claims related to critical diseases. This finding also suggests that insurers find claims in C5 and C3 to be more suspicious, as such diseases require extensive medical treatment. If this is the case, then insurers are more likely to dedicate more time investigating such claims to verify that the provided treatment and medications match the disease type.

HCPs are more involved in the healthcare system because they provide medical services and file medical claims directly to insurance companies; therefore, we argue that our control variables are likely to influence the likelihood of committing fraud and the probability of rejecting the claim.[65]

$$P(\text{Fraudulent\_Claims}_{it}) = \alpha + \beta_1 \sum \text{Fraudester}_{it} + \gamma X_i + \varepsilon_{it} \tag{3.1}$$

where Fraudulent_Claims is a binary variable with a value of 1 if the claim is fraudulent and rejected (reported by the insurer) and zero otherwise. As we mentioned previously, we utilize the insurer's fraud report and the investigator's explanation for all fraud schemes to create several measures of our key variable, namely, "fraudsters." X_i is a vector of our control variables (see Table 3.3 for variable definitions).

[64] Crocker and Tennyson (2002) state that insurance companies mitigate fraud and suspicious claims by separating bad claims from good ones through auditing by utilizing observable claim/insured characteristics for the company.
[65] Cummins and Tennyson (1996) argue that including such control variables would also separate the effects of moral hazard on claiming behavior from the effects on fraud behavior.

Table 3.4 reports the summary statistics of our dataset. The table shows that fraudulent claims from HCPs represent 76.2 percent of all fraudulent claims. However, fraudulent claims from PHs represent 16.5 percent, and the remainder are fraudulent claims by the two parties in concert, representing 7.3 percent. These univariate statistics provide some support for H1, which states that HCPs are more likely to commit fraud/abuse in the healthcare insurance market. Regarding our control variables, the table reveals that the average size of all claims is approximately USD $17.71, the average rejected amount is USD $2.3, the average settlement period is 19 days, and the average filing period is 20 days.

Regarding the PH medical visit characteristics, Table 3.4 shows that most medical visits were to emergency centers (16.6%), labs (14.6%) and to general medical doctors (14.4%). The HCPs that provide heavy and intense medical treatments comprise 4.1 percent (inpatient facilities) and 2.2 percent (medical service centers and radiology centers) of the medical visits. Furthermore, most of the claims were filed by PHs who hold the least generous coverage types (i.e., first and second medical classes), and only 3.9 percent were under the most generous coverage type (i.e., suite and executive). Regarding disease type (ICD9 chapters), 34.9 percent of claims were filed by PHs who suffered from diseases of the respiratory system, and the remainder suffered from the following: symptoms, signs, and ill-defined conditions (12 percent); injury and poisoning (11.89 percent); and dermatologic, musculoskeletal and rheumatologic disorders (10.57 percent). Some PHs suffered from more critical diseases (i.e., endocrine, nutritional and metabolic diseases and immunity disorders), and their claims represent 4.37 percent. Finally, most PHs were treated within the HCP network (99%), and the remainder were treated outside the medical network.[66]

Finally, regarding PH characteristics, Table 3.4 shows that 59.5 percent of the claims were filed by males, 40 percent were filed by married PHs, and the average age of all PHs was 29.5 years. Finally, regarding PH dependents, most claims were filed by the head of the family (HOF), 16.8 percent were filed by spouses, and the remaining 22.1 percent were filed by children.

Table 3.4. Summary Statistics

Column	1	2	3	4	5
	N	Mean	Std. Dev.	Min	Max

[66] The average claim size for claims outside the HCP network was USD $127.79, and the average claim size for claims within the network was USD $18.2.

VARIABLE					
Claim Characteristics					
Claim_Size	605037	13.37	48.25	0.2	6992.18
Due_Value	605037	10.28	36.55	0	6392.18
Settlement_Days	605037	19.59	12.54	0	381
File_Days	605037	20.01	12.73	0	738
Coinsurance	605037	0.79	4.72	0	1735.2
Rejected	605037	2.30	18.34	0	4238.73
No_Network	605037	0.00	0.06	0	1
Fraud Side (Fraudsters)					
PH_Fraud	10236	0.16	0.37	0	1
HCP_PHFraud	10236	0.07	0.26	0	1
Process Type (Claim Disposition)					
Fully_Covered	605037	0.20	0.40	0	1
Fraudulent_Claims	605037	0.01	0.12	0	1
Healthcare Provider (HCP)					
Doctor	605037	0.14	0.35	0	1
Emergency Center	605037	0.16	0.37	0	1
Inpatient Facility	605037	0.04	0.19	0	1
Lab	605037	0.14	0.35	0	1
Radiology Center	605037	0.02	14.98	0	1
Medical Service Center	605037	0.00	0.02	0	1
Medical Coverage Type					
Suite/Executive	605037	0.03	0.19	0	1
Second	605037	0.03	0.18	0	1
Third	605037	0.07	0.26	0	1
Dependents					
Spouse	605037	0.16	0.37	0	1
Child	605037	0.22	0.41	0	1
ICD9 Chapters					
ICD9_Cat2	605037	0.00	0.08	0	1
ICD9_Cat3	605037	0.10	0.30	0	1
ICD9_Cat4	605037	0.12	0.32	0	1
ICD9_Cat5	605037	0.11	0.32	0	1
ICD9_Cat6	605037	0.01	0.12	0	1
ICD9_Cat7	605037	0.04	0.20	0	1
ICD9_Cat8	605037	0.04	0.20	0	1
ICD9_Cat9	605037	0.04	0.20	0	1
ICD9_Cat10	605037	0.08	0.27	0	1
ICD9_Cat11	605037	0.05	0.22	0	1
ICD9_Cat12	605037	0.01	0.11	0	1
Age	605,037	29.48	15.40	0.035	117
Male	605037	0.59	0.49	0	1
Married	605037	0.40	0.49	0	1

3.4 Probit Regression Results

Table 3.5 shows the probit regression results of fraudulent rejected claims on the fraudster side (estimating Equation 1). The dependent variable in Column (1) is fraudulent claims and is defined as a binary variable with a value of one if the claim is reported as fraudulent and zero otherwise. The key variables of interest are H*CP_Fraud, PH_Fraud,* and *HCP_PHFraud*. HCP_Fraud is defined as a binary variable with a value of one if fraud is committed from the HCP side (the admitted variable) and zero otherwise. PH_Fraud is a binary variable with a value of one if fraud is committed from the PH side and zero otherwise. HCP_PHFraud is a binary variable with a value of one if fraud is committed due to collusion by PHs and HCPs and zero otherwise.

When we examine the estimated coefficients of the fraudster side variables (H*CP_Fraud*, *PH_Fraud*, and *HCP_PHFraud)*, we find that fraud is most likely to be committed from the HCP side. We also find that claims are more likely to be rejected due to fraudulent actions from HCPs. For instance, fraud/abuse from the PH side is approximately 40 percent lower than fraud from the HCP side (the omitted variable). This finding is consistent with our expectations, as a PH cannot file a claim without action from an HCP and has less understanding of how the claiming system works. Furthermore, to a large extent, PHs are less likely to be familiar with the provided medical services and medications, which gives HCPs almost full control of claims. In addition, HCPs are capable of abusing/misusing insurance by providing, for instance, unnecessary medical services to gain more revenues from insurance. All of these factors place HCPs in an advantageous position for defrauding PHs and insurance companies. Thus, our results support H1. However, there are cases where PHs and HCPs collude to defraud insurance companies. The results show that fraud as a result of collusion between the two sides is 25 percent lower than fraud from the HCP side.

Table 3.5. Probit Regression Results: Fraudulent Claims and Fraudster Sides

The table shows the probit regression results of fraudulent rejected claims on the fraudster side. The coefficients of this probit model are converted into a marginal value. The dependent variable in Column (1) is the fraudulent claims and is a binary variable with a value of one if the claim is reported as fraudulent and zero otherwise. The key variables of interest are HCP_Fraud, PH_Fraud, and HCP_PHFraud. HCP_Fraud is a binary variable with a value of one if fraud is committed from the healthcare provider (HCP) side *(the admitted variable)* and zero otherwise. PH_Fraud is a binary variable with a value of one if fraud is committed from the policyholder (PH) side and zero otherwise. HCP_PHFraud is a binary variable with a value of one if fraud is committed due to collusion by PHs and HCPs and zero otherwise. For all other control variable definitions, refer to Table 3-3. Robust standard errors are reported in parentheses, and *, **, and *** indicate significance at the 10%, 5% and 0.1% levels, respectively.

Variable	(1) Fraudulent Claims
PH_Fraud	-0.407***
	(0.0076)
HCP_PHFraud	-0.254***
	(0.0161)
Claim_Size	-0.001***
	(0.0002)
Spouse	0.027**
	(0.0136)
Child	-0.122***
	(0.0147)
Male	0.055***
	(0.0099)
Age	-0.002***
	(0.0004)
Married	-0.141***
	(0.0118)
Doctor	0.198***
	(0.0120)
Emergency Center	0.200***
	(0.0156)
Inpatient Facility	0.037*
	(0.0228)
Lab	0.101***
	(0.0128)
Radiology Center	0.137***
	(0.0286)
Medical Service Center	-0.084
	(0. 1096)
Suite_Class	0.548***
	(0.0977)
Executive	-0.115***
	(0.0197)
Second_Class	-0.056**
	(0.0229)
Third_Class	-0.252***
	(0.0112)
ICD9 Dummy Variables	Yes
Observations	10,236
Pseudo R2	0.22
pseudolikelihood	-5310.86

However, one might argue that the results shown in Table 3.5 are not robust, as the classification of fraudulent claims from the insurer side might be biased and that some of these rejected claims are likely to be valid. For instance, some claims are rejected due to late filing (e.g., a claim should arrive within 45 days of the service date) or because the policy does not cover the disease. The argument is that the HCP is not expected to know all of the diseases and medications covered by the policy and that the delay in filing claims might be due to time constraints. An additional argument is that when insurance companies audit claims and reject some, this might not only indicate potential fraud but also provide evidence of mistakes in filing claims.

Table 3.5 shows that most of our control variables are significant at the 1 percent significance level. When we examine the estimated coefficients for medical coverage types, we find that the estimated coefficients for the suite are positive and statistically significant at the 1 percent level. The suite claims are 54.8 percent more likely to be fraudulent than the first-class claims (the omitted variable). This finding suggests that claims from the suite medical class are likely to be fraudulent, and the probability of rejecting claims from such classes is larger than that of rejecting claims from the first class. One possible explanation is that suite medical classes provide luxury services and that HCPs may abuse insurance by providing unnecessary services and medications. Thus, suite claims are more extensive than first-class coverage claims. These extensive claims could, in fact, cause insurers to audit such claims to ensure that the correct services and medications are provided. Thus, the extent of these claims makes them more likely to be rejected due to the likelihood of fraudulent actions by the HCPs. However, the coefficients of the second, third, and executive medical classes are negative and significant at the 1 percent level, suggesting that these claims are less likely to be fraudulent and to be rejected than claims of the first class. Claim size and settlement period are negative and significant at the 1 percent level. These results indicate that small claims with a short settlement period are less likely to be fraudulent and to be rejected than large claims with longer settlement periods.

Interestingly, the coefficient of the male variable is positive and significant at the 1 percent level, suggesting that claims from males are 5.5 percent more likely to be fraudulent than those from females. The coefficient for married is negative and significant at the 1 percent level, indicating that claims from married PHs are 14 percent more likely to be fraudulent and rejected than those from single PHs. Almost all the HCP providers' coefficients are positive and significant at the 1 percent level. For instance, claims from doctors and emergency centers are 20 percent more likely to be fraudulent than those from the omitted group, pharmacies. When we examine the estimated coefficients of the dependent variables, we find that the child variable is negative and significant. This finding indicates that a child's claims are less likely to be fraudulent and rejected than claims from the HOF, the omitted group. However, the coefficient of the spouse is positive and significant at the 10 percent level, as spouse claims are 2.7 percent more likely to be fraudulent than those from the omitted group, HOF. The coefficient of age is negative and significant at the 1 percent level, indicating that claims from older PHs are less likely to be fraudulent.

To rule out the potential effect of any filing claim mistakes on healthcare insurance fraud, we repeat our analysis by including only the most fraudulent claims—given the insurer's fraud report—and then compare the results with the full fraudulent claims sample. Our subgroup analysis includes seven processing codes from the HCP side (H02, I03, T10, T14, T21, O03, and V00) and three codes from the PH side (E12, L02, and l03). Table 2 reports the frequency of those claims, along with the processing codes, processing results, and the investigator's explanation for all fraud schemes.

The new results are reported in Table 3.5A and remain consistent with our expectations: ruling out mistakes in filing claims or the misclassification of fraudulent claims from the insurer side, the individual PH's fraudulent behavior was 12 percent lower than that of the HCP.

Table 3.5A Probit Regression Results: Robustness Test for Fraudulent Claims

The table shows the probit regression results of fraudulent rejected claims on the fraudster side. The coefficients of this probit model are converted into a marginal value. The dependent variable in Column (1) is the fraudulent claims and is a binary variable with a value of one if the claim is reported as fraudulent and zero otherwise. The key variables of interest are HCP_Fraud, PH_Fraud, and HCP_PHFraud. HCP_Fraud is a binary variable with a value of one if fraud is committed from the healthcare provider (HCP) side *(the admitted variable)* and zero otherwise. PH_Fraud is a binary variable with a value of one if fraud is committed from the policyholder (PH) side and zero otherwise. HCP_PHFraud is a binary variable with a value of one if fraud is committed due to collusion by PHs and HCPs and zero otherwise. For all other control variable definitions, refer to Table 3-3. Robust standard errors are reported in parentheses, and *, **, and *** indicate significance at the 10%, 5% and 0.1% levels, respectively.

Variable	(1) Fraudulent Claims
PH_Fraud	-0.12***
	(0.0076)
HCP_PHFraud	-0.14***
	(0.0161)
Control Variables	Yes
Observations	5,493
Pseudo R2	0.38
Pseudo Likelihood	-1510.82

3.5 Data and Methodology: PRIDIT Model Analysis

3.5.1 Methodology: PRIDIT Model

To further investigate healthcare insurance fraud detection by using an unsupervised learning method for fraud detection rather than the supervised models (e.g., logistic regression) used in the previous section, we apply principal component analysis of the RIDIT scores (PRIDIT) in this section. Insurance economists have used PRIDIT for fraud detection in automobile insurance (e.g., Brockett et al. 2002, Ai et al. 2013, and Golden et al. 2019) and in the assessment of hospital quality (e.g., Lieberthal and Comer 2013, Lieberthal 2008). PRIDIT obtains fraud suspicion scores for claims and without relying on and/or accessing audited claims, it classifies them by creating fraud predictor variables.

PRIDIT assigns each claim a relative position of fraud suspicion, and the predictor variables then generate information about each suspicion position. For instance, the medical network can be a predictor variable that red-flags claims based on medical visits and proximity to medical doctors. Thus, investigators and adjusters are likely to question claims from patients who visited medical facilities outside the medical network. Additionally, investigators might question the intention of patients who drive long distances to visit medical doctors, especially when nearby medical facilities are available to provide the same medical treatments.

PRIDIT separates claims into two classes (positive vs. negative scores) and determines the relative degree of confidence in that fraud classification. However, interpreting the results can be confusing; we aim to explain how it works with the example below. Table A1 presents the following four predicting variables, namely, (a) gender, (b) dependents, (c) marital status, and (d) age, along with the corresponding PRIDIT weight, RIDIT score and PRIDIT score.

1. PRIDIT is the sum of an individual's variable scores (i.e., -0.20 in our example) over all predicting variables. The claims with negative scores are considered part of the "suspicious" class, while those with positive scores are considered part of the "nonsuspicious" class (see Column 3 in Table A1, for example). The larger the magnitude of these PRIDIT scores is, the stronger the membership in the class (Brockett et al. 2002). For example, the age variable with a PRIDIT score of 0.34 is slightly less suspicious than the dependent variable with a PRIDIT score of 0.26.
2. Following Brockett et al. (2002), the individual predictor score for each claim (RIDIT score) equals the subtraction of the proportion of claims below each category (e.g., category 1 for males in Table A2) and the proportion of claims above the same category (in this case, category 2 for

females in Table A2). Table A2 presents an example of the RIDIT score for gender. For instance, assume that 75 percent of the claims are filed by "males" and 25 percent are filed by "females." In this case, the proportion of claims below category 1 (males) would be zero (see Column 4 in Table A2), and the proportion of claims above category 1 would be 75 percent (see Column 5 in Table A2). Thus, the RIDIT scores for category 1 (males) and category 2 (females) would be -0.25 and 0.75, respectively (see Column 6 in Table A2).

3. The PRIDIT model also provides an optimal weighting used to determine the PRIDIT weights, where each variable has a weight from -1 to 1 (see Column 1 in Table A1). The variables with a larger weight dominate the PRIDIT model and drive the suspiciousness of the medical claims. For instance, positively weighted variables are positively associated with suspiciousness, while negatively weighted variables are inversely associated with suspiciousness.
   a) The PRIDIT weights are a product of the correlation between the individual predictor score for each claim (RIDIT score) and the overall summative PRIDIT score for the same claim.
   b) These PRIDIT weights provide a "variable discriminatory power weight" for the individual predictor variable. The variables that are more highly correlated with the overall summative score will have a higher weight (Brockett et al. 2002, Ai et al. 2003, and Golden et al. 2019).
   c) The correlation can be positive or negative (ranging from -1 to 1). The positive correlation with PRIDIT scores comes from the consistent ordering of the categories based on the ordinal predictor variables (see Table 3.6: PRIDIT variable list) such that the lower end of the variable categories is more suspicious in relation to fraud than the upper end (e.g., category 1 for males is more suspicious than category 2 for females). This is expected to be true for all variables by construction. When we sum an individual's variable scores over other variables, then the lower summative overall PRIDIT score (i.e., -0.20 in our example) corresponds to a more suspicious claim file.
   d) However, there may be a small number of variables with negative weights, meaning that the individual predictor variable correlates negatively with the overall summative claim file score (see the marital status variable, Column 1 in Table A1). A negative weight means that while the individual variable still predicts, it does so in a reverse direction. Overall, predictor variables are likely to relate positively to the overall score, and the extent of the association is provided by the variable weight (Brockett et al. 2002, Ai et al. 2003, and Golden et al. 2019).[67]

[67] For more information about the PRIDIT model, refer to Brockett et al. (2002), Ai et al. (2013), and Golden et al. (2019).

e) Interpreting negative weights for categorical variables can be confusing depending on how the variables are ordered by rank (recall Step b). In this study, we follow Brockett et al. (2002) in ordering our variables so that the lowest value is associated with the highest probability of fraud.

f) For instance, our dependents variable has three categories (1 for principal HOF, 2 for spouse, and 3 for child). Given this ranking, if the weight is positive, we presume that claims from principals are more suspicious than claims from a child.

g) However, if this predictor variable has a negative weight (e.g., -0.5), then a higher category (not a lower category) is more indicative of fraud. As the category increases (e.g., from 1 to 2), the overall score decreases (could go below zero and become more suspicious). This negative weight makes a "principal" first-category rank indicative of nonfraud (because of the exhibited negative relationship with overall fraud suspiciousness), and the contribution of a first-category rank should increase the overall assessment of non-fraud.

h) Overall, regardless of whether the weights are positive or negative, they help insurance companies to identify the variables that drive the overall PRIDIT score and then to understand the extent of the association with them (see Brockett et al. 2002, Ai et al. 2003, and Golden et al. 2019).

Overall, PRIDIT, as an unsupervised model, is more useful in fraud detection than supervised models (e.g., logistic regression). Dionne et al. (2009) state that all supervised models are likely to rely on biased samples to estimate parameters. Furthermore, the auditing of claims can cost much money as fraud patterns evolve and fraudsters learn how to avoid getting caught. Thus, using unsupervised methods to detect fraud is very cost-efficient. Ai et al. (2009) state that the model is useful because all predictor variables (categorical, binary, and continuous) can be utilized. Furthermore, PRIDIT provides hints about fraudulent patterns by employing predictor variables to quantify suspicious/fraudulent activities. It demonstrates the importance of predictor variables in quantifying suspicious/fraudulent activities as cases of "claims suspicion."

3.5.2 ***Data***

We recode the data previously described in Table 3.3 (Section 3.3.1) to conform to the underlying assumptions of the PRIDIT method. PRIDIT was applied to the following variables listed in Table 3.6: the demographic characteristics of the claimant, the financial characteristics of the claim, the insurance variables for the claimant, and the clinical characteristics of the claim. We also use clinical characteristics to divide the data into inpatient and outpatient claims. As a result, the provider type is not used for inpatient claims since all inpatient claims are coded as being of the same provider type, the hospital. To apply the PRIDIT model, we need to rank-order our categorical variables based on the risk of "potential fraud" (see Brockett et al. 2002, Ai et al. 2003, and Golden et al. 2019). Table 3.6 shows that our variables are ordered in such a way that the lowest value is associated with the highest probability of fraud (Brockett et al. 2002).

Table 3.6. PRIDIT Variable List

This table shows a list of variables with descriptions and their rank of suspiciousness if the variable is categorical. We use these variables in PRIDIT. Our categorical variables are ordered in such a way that the lowest value is associated with the highest probability of fraud (see Brockett et al. 2002, Ai et al. 2013, and Golden et al. 2019).

Variable	**Type**	**Description of Variables and their PRIDIT Rank Order**
Religion	Categorical	1 if Muslim, and 2 if Non-Muslim
Insurance Type	Categorical	I if commercial insurance and 2 if Takaful insurance
Culture[68]	Categorical	1 if developing country, and 2 if developed country.
Claim Size	Continuous	Total claimed amount filed by the policyholder (PH) (size of claim)
Paid Amount	Continuous	Total amount paid to the PH
Dependents	Categorical	1 if the principal of the policy, 2 if the spouse, 3 if a child, and 4 if other (e.g., parent or sibling)
Gender	Categorical	1 if male and 2 if female
Marital Status	Categorical	1 if married and 2 if single
Age	Continuous	Age of the PH; the age group (31-45) is the most suspicious group
Length of Stay in Medical Facility	Continuous	Length of stay (days) in medical facility
Copayment Amount	Continuous	Out of pocket paid by the PH
Beneficiary Status	Categorical	Status of the insurance policy: 1 if active and 2 if deleted.
Chronic	Categorical	1 if the PH has a chronic condition and 2 if not

[68] Prior insurance literature relies on samples of developing versus developed countries to examine the relationship between insurance demand and culture (e.g., Religion). For more information, see Outreville (2018).

Claim Type	Categorical	How claims are filed: 1 if a reimbursement, 2 if through the provider portal (preapproval), 3 if back office, 4 if an electronic claim, and 5 if filed via a medical network
Coverage Type	Categorical	Type of insurance coverage: 1 if a worker plan, 2 if a basic coverage plan, 3 if an economy plan, 4 if a regional plan, and 5 if a VIP plan
Policy Length	Continuous	Length of policy (days)
Provider Type	Categorical	Type of medical facility: 1 if a pharmacy, 2 if a hospital, and 3 if a diagnostic/polyclinic center
Medical Visit Type	Categorical	Type of medical visit: 1 if dental, 2 if maternity, 3 if inpatient, and 4 if outpatient

3.5.2 Summary Statistics and Weights

PRIDIT is a fully nonparametric method that is designed to produce a score with a mean of zero. However, the distribution of the data will drive all other moments and the shape of the distribution. We present the results below for the summary statistics of the data in Table 3.7. These results show that the means of all our variables are zero, indicating that PRIDIT was correctly applied.[69] Furthermore, there was a difference in the distribution of scores when dividing our data into inpatient data only, outpatient data only, Takaful claims only, and commercial claims.

Next, we describe the meaning of our PRIDIT results for each analysis (all data, inpatient data only, outpatient data only, Takaful claims only, and commercial claims).

[69] This means that the average claim is classified as neither suspicious nor nonsuspicious. Claims with positive scores are classified into the nonsuspicion class, whereas claims with negative scores are placed in the suspicion class.

Table 3.7. Summary Statistics for PRIDIT Scores

Summary Statistic	Mean*	STD	10%	25%	50%	75%	90%	Obs.
All Data	0.000	1.55	-1.58	-1.13	-0.31	1.04	1.93	633,042
Commercial Scores Only	0.000	1.87	-0.26	-0.21	-0.14	-0.04	0.14	68,808
Takaful Scores Only	0.000	1.57	-1.67	-1.23	-0.41	1.17	2.16	564,234
Muslim Scores Only	0.000	1.59	-1.55	-1.12	-0.39	1.08	1.92	475,531
Non-Muslim Scores Only	0.000	1.53	-0.69	-0.54	-0.26	0.16	0.79	157,511
Inpatient Scores Only	0.000	1.59	-1.40	-1.04	-0.19	0.89	1.42	59,57
Outpatient Scores Only	0.000	1.54	-1.78	-1.28	-0.41	1.01	2.50	584,239

*Values shown as “0.000” are less than 0.001 in magnitude.

3.6 PRIDIT Results

We start by applying the PRIDIT analysis to all claims to determine how useful the model is in determining the relative suspicion of claims. The ultimate goal is to identify the variables and claims that have the strongest indicators of suspicion. To do so, we also apply the PRIDIT model to insurance type (Takaful (Islamic) versus commercial insurance) and medical visit type (inpatient versus outpatient).[70]

We argue that given the nature of healthcare insurance, fraud patterns are not significantly different across healthcare insurance markets, whether commercial, Takaful, public, or private healthcare insurance markets.

The histograms of the PRIDIT scores for all groups show that the vast majority of claims are clustered in the positive area of the distribution, representing non-suspicious claims.[71] However, across all groups, a thin left-tailed distribution is also reflected in the skew of the distribution.[72] The tail represents the highly negatively weighted scores for the relatively small number of claims that have high suspicion scores (scores above -1 or -2), which are an excellent target for further analysis by trained fraud detection investigators and other review methods. Thus, in a totally objective and automated manner, an insurance company can decide to investigate claims with negatively weighted scores (higher than -1 or -2) and pay

[70] Note that most published fraud detection papers focus on the medical visit type (e.g., Parente, et al., 2012). For the sake of this dissertation, we focus on various filing types. However, for publication purposes, we need to focus our analysis on one filing type rather than on various filing types (religion, nationality, provider, etc.). For more information, see Ai, et al. (2018).
[71] The histogram of the PRIDIT score results is not reported in this paper and is available upon request
[72] The t-statistics between the subsample groups are nonsignificant for all scores, scores above 1 (non-suspicious), and claims below -1 (very suspicious). The results are not reported here but are available upon request.

out the remaining claims (or, if resources allow, investigate claims in increasing PRIDIT score order until resources are exhausted).

Table 3.8 shows the PRIDIT weights for our variable list presented in Table 3.6 across all groups (all data, inpatient data only, outpatient data only, Takaful claims only, and commercial claims). The weights help to identify the main variables that drive the PRIDIT scores and/or dominate the PRIDIT model. The positively weighted variables are positively associated with to refine our results further, while the negatively weighted variables are inversely associated with suspiciousness (for more information, see Section 2.5.1).

The first important result to report is consistent with our expectations in which, given the healthcare providers involvement in the healthcare insurance market, our key variables: policyholders' religion culture (developed versus developing countries), and insurance type (Takaful versus commercial insurance) have small weights across all subgroups. This indicates that our key variables are not strong in driving the PRIDIT scores, and are not crucial when it comes to identify fraud in the healthcare insurance market (providing support for H3, and H4). We argue that policyholders' culture, and religious affiliations may be crucial factors on influencing abusive behaviors. However, given healthcare providers' involvement with the medical process, including claims process, it would be complicated to capture the overall effect of their culture, and religious affiliations on defrauding insurance companies.

Table 3.8 shows that no single variable is a strong predictor for suspicion in and of itself. Certain variables are more predictive of suspicion—claim size, total amount paid, gender, marital status, age group, and dependents (HOF, spouse, or child) are positively weighted, dominating the model across all subgroups and driving the overall PRIDIT score. The reported results suggest that for categorical variables, as the category increases (from 1 to 2), the overall PRIDIT score increases (making it less suspicious), indicating thus that claims from principals are more suspicious than claims from a child and helps claims auditors investigate claims efficiently.[73]

Large claims—hypothetically—seem more suspicious and therefore are more likely to be at least partially rejected. However, one might argue that smaller claims are relatively more suspicious; to avoid

[73] HOF "principals" are the ones who buy the insurance and provide medical coverage to family members if married. Thus, the likelihood of misusing the medical insurance and filing a fraudulent claim is greater for them than for a spouse or child, assuming that the principal is an opportunistic person. Furthermore, most claims are filed by the HOF. Thus, when we assume that the PH is honest and the HCP is opportunistic, we should expect more fraudulent claims from HOFs given HCP involvement.

getting caught, a fraudster might submit a series of smaller claims to avoid detection.[74] This makes sense, as fraud patterns are evolving. Our analysis shows that more small claims belong to the suspicion class and have a negative score (as the claim size increases, the overall PRIDIT score increases, making it less suspicious). One other important variable is the beneficiary age, which is also positively weighted, indicating that claims from older beneficiaries (aged over 60) are relatively less suspicious than claims from younger PHs (category 1: age group 31-45). Our results indicate that claims fraud detection could benefit from the additional analysis of the characteristics of the patients.

Table 3-8. PRIDIT Variable Weights

Table 3.8 shows the PRIDIT weights for our variable list presented in Table 3.6 for the following five groups: (1) all data, (2) Takaful claims only, (3) commercial claims only, (4) outpatient data only, and (5) inpatient data only. The weights help to identify the main variables that drive the PRIDIT scores and/or dominate the PRIDIT model. The positively weighted variables are positively associated with suspiciousness, while the negatively weighted variables are inversely associated with suspiciousness.[75]

Column	(1)	(2)	(3)	(4)	(5)
Variable	All Data	Takaful Weights	Commercial Weights	Outpatient Weights	Inpatient Weights
Religion	0.08	0.10	-0.01	0.04	-0.05
Insurance Type	-0.18	-	-	-0.23	-0.29
Claim Size	0.32	0.28	0.50	0.04	0.38
Paid Amount	0.35	0.27	0.51	0.06	0.45
Dependents	0.46	0.50	0.01	0.58	0.37
Gender	0.33	0.37	0.02	0.37	0.23
Marital Status	0.29	0.30	-0.02	0.43	0.25
Age	0.32	0.34	0.00	0.45	0.30
Length of Stay in Medical Facility	0.16	0.13	0.48	0.00	0.27
Copayment Amount	0.26	0.09	0.47	-0.01	0.38
Beneficiary Status	-0.02	0.00	0.00	-0.02	-0.01
Chronic	0.08	0.06	0.00	0.08	0.00
Claim Type	-0.13	-0.13	-0.03	-0.11	-0.16
Coverage Type	0.30	0.36	0.03	0.32	0.24

[74] Crocker and Tennyson (2002) argue that "…small amounts of fraud tend to be relatively easy to accomplish, while more substantial attempts are more costly to implement successfully."
[75] For more information, see Section 1.5.1 Methodology: PRIDIT Model and refer to the following three papers: Brockett et al. (2002), Ai et al. (2013), and Golden et al. (2019).

Policy Length	0.00	0.01	-0.01	0.00	0.01
Provider Type	0.03	0.04	0.02	-0.02	0.05
Medical Visit Type	-0.21	-0.24	-0.17	-	-

Interestingly, across all groups, the insurance type variable (Takaful versus commercial) is negatively weighted, indicating that the type of insurance market can be relatively suspicious. However, the direction of suspicion is reversed over the categories, with a higher category (rather than a lower category) being more indicative of fraud.[76] In our case, for instance, commercial insurance is a lower category; thus, its claims are less indicative of fraud than Takaful insurance claims. One possible explanation is the heterogeneity of these insurance market business models (for more information, see Section 1.2, the Takaful insurance section, in essay one). Another possible explanation is the heterogeneity of the PH culture and the religious affiliation between the Takaful and commercial insurance markets, which may make it difficult to determine the independent effect of each variable on suspicion. The results thus suggest that we may need to further divide our analysis by Hofstede cultural variables (e.g., individualism, power distance, masculinity, and uncertainty) and relate them to healthcare insurance fraud to refine our results further. We plan to conduct this additional analysis, which is beyond the scope of this thesis, in future studies.

An additional variable that is negatively weighted across most groups (ranging from -0.17 to -0.24) is the medical visit type, indicating that this type is relatively suspicious. The results are consistent with our expectations that the degree of suspicion is lower in the inpatient claims group than in the outpatient claims group. For instance, in our case, inpatient visits are a lower category, and thus, claims for that type are less indicative of fraud than claims for outpatient visits (supporting H2). One reason may be the heterogeneity of outpatient claims, which may need to be further divided by service type or clinical area. These results are also consistent with the results of analyses in the US, which have shown a low rate of fraud in inpatient claims (e.g., Parente et al., 2012). We argue that many fraud patterns, beyond the basic patterns, can be detected only when tracking a series of events. For example, in inpatient medical visits, surgery itself is likely to be legitimate. However, the fact that a patient was given certain drugs a few days or a few weeks after the surgery might be suspicious if the drug is not necessary.[77]

[76] Our categorical variables are ordered so that the lowest value is associated with the highest probability of fraud. For more information, see Table 2-6: PRIDIT Variable List.

[77] One benefit of having a large dataset and applying the PRIDIT model is that we do not necessarily have to know which medical service is valid or what drugs are not associated with what surgeries; we can simply allow the unusual patterns to raise red flags in the PRIDIT model.

3.7 Classifying Claims by the PRIDIT Score: Which Claims Should Be Audited?

When insurers receive claims, they spend time auditing them until they are settled. They verify claims through investigation or auditing to ensure that they are valid. During the verification process, they utilize all observable claim/PH characteristics (e.g., claim size, provider type, medical coverage type) and certain other factors, including the type of medical service.[78] They spend more or less time, depending on the case. If they suspect that some claims are invalid or suspicious, they are likely to spend more time investigating them. A longer settlement time in such cases enables insurers to obtain more information about the claims and therefore provide a fair offer to compensate the claimant for the actual economic loss. However, the methods of committing insurance fraud continue to evolve and thus require ongoing identification efforts, which can be costly. These costs are typically passed on to PHs in the form of higher premiums. Thus, identifying the most suspicious claims would allow insurance companies to handle claims efficiently and reduce the overall fraud rate.[79]

After generating the overall PRIDIT score for all claims (Section 3.6), the model classifies them into suspicion classes according to the underlying negative or positive overall score. Brockett et al. (2002) utilize 127 claims and implement a two-way classification in which claims with negative scores are placed in class 1, "potential fraudulent claims," and claims with positive scores are placed in class 2, "unsuspicious claims." We follow them in implementing four-way classifications to order the classes so that the lowest value is associated with the highest probability of suspicion. We create the following four classes:

1. Class 1 includes claims with scores lower than -2 ($-2 > x$, where x is the overall PRIDIT score): such claims are likely to be within the thin left-tailed distribution reflected in the skewness of the distribution, and they constitute a relatively small number of claims with high-suspicion scores.
2. Class 2 includes claims with scores greater than or equal to -2 and lower than -1 ($-2 \leq x < -1$): such claims belong to the moderate suspicion class.
3. Class 3 includes claims with scores greater than or equal to -1 and lower than 0 ($-1 \leq x < 0$): such claims belong to the low-suspicion class.
4. Class 4 includes claims with scores greater than or equal to zero ($0 \leq x$, where x is the overall PRIDIT score): such claims belong to the nonsuspicion class.

[78] Crocker and Tennyson (2002) state that insurance companies mitigate fraud and suspicious claims by separating bad claims from good ones through auditing by the utilization of the observable claim/insured characteristics for the company.
[79] Butler and Francis (2010) state that if a claim is handled appropriately, PHs are likely to stay longer, potentially increasing the overall retention rate. If claims are not handled properly, the insurer is more likely to lose customers as well as damage its overall reputation.

Table 3.9 shows a random sample of 14 claims, their transformed RIDIT value, their overall PRIDIT score, and their suspicion class. The claims are ranked in descending order, where class 4 is the nonsuspicion class and class 1 is the highest suspicion class. Applying PRIDIT allows claims to be audited efficiently. Instead of auditing claims randomly or based on a subjective assessment, an insurance company, given the PRIDIT results, may investigate claims 12, 13, and 14 when the model provides red flags about them. If more resources are available, they can then audit claims 9, 10, and 11. The insurance company will then keep investigating claims in increasing class order (from class 1 to class 4) until its resources are exhausted (Brockett et al. 2002). This procedure would allow it to efficiently investigate any claims that belong to the suspicion class. Brockett et al. (2002) state that by doing so, insurance companies can achieve internal consistency and experience significant economic savings. Furthermore, the PRIDIT scores and suspicion class rankings can help third parties (e.g., investigators and adjusters) focus on specific variables and claims because the PRIDIT scores provide these assessors with a mathematical backup for their subjective assessment, thus helping them perform their job.

Table 3.9. PRIDIT Scores and their Suspicion Classes for a Sample of Claims.

The table shows a sample of claims, their transformed variable values, their overall score, and their suspicion class. The variables in Column (8) are the overall PRIDIT score. The key variable of interest is the suspicion class in Column (9). Class 1 is the highest suspicion class, while Class 4 is the nonsuspicion class. The classes are ordered so that the lowest value is associated with the highest probability of suspicion.

Column	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Claim Number	Dependents	Gender	Marital Status	Claim Size	Amount Paid	Scores	Suspicion Class
1	-0.75	-0.73	-0.66	0.14	0.01	0.26	4
2	-0.75	-0.73	-0.66	0.82	-0.12	0.23	4
3	-0.75	-0.73	-0.66	-0.08	-0.07	-0.50	3
4	-0.75	-0.73	-0.66	-0.16	-0.10	-0.64	3
5	-0.75	-0.73	-0.66	0.05	-0.12	-0.01	3
6	-0.75	-0.73	-0.66	0.12	0.03	-0.40	3
7	-0.75	-0.73	-0.66	-0.04	-0.05	-0.47	3
8	-0.75	-0.73	-0.66	-0.13	-0.09	-0.27	3
9	-0.75	-0.73	-0.66	-0.17	-0.10	-1.25	2

10	-0.75	-0.73	-0.66	-0.08	-0.06	-1.59	2
11	-0.75	-0.73	-0.66	-0.15	-0.10	-1.33	2
12	-0.75	-0.73	-0.66	-0.17	-0.11	-2.10	1
13	-0.75	-0.73	-0.66	-0.20	-0.12	-2.09	1
14	-0.75	-0.73	-0.66	-0.12	-0.09	-2.06	1

3.8 Conclusion

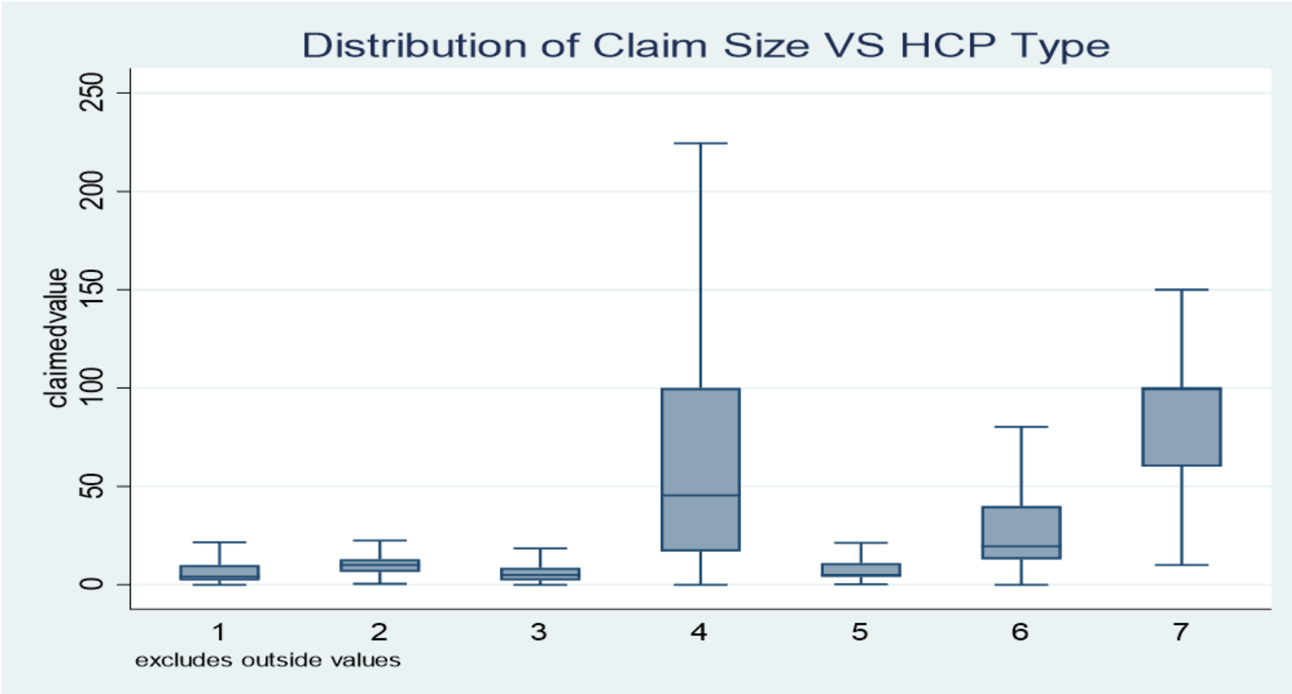
We first compare and contrast potential fraud schemes in private healthcare insurance markets. We then examine whether and to what extent fraud is committed largely by HCPs rather than healthcare insurance PHs. Finally, we apply principal component analysis of RIDIT scores (PRIDIT) to further deepen our understanding of healthcare insurance fraud and to identify the variables and claims with the strongest indicators of suspicion.

Based on a probit model, our results show that fraud/abuse activities in the healthcare insurance market are committed mostly by HCPs rather than by PHs. Furthermore, fraudulent claims are more likely to be rejected when the HCP rather than the PH controls the claim (supporting H1). This result is consistent with our expectations that given the HCP's involvement with the medical process, including the claims process, it is difficult for the PHs to commit fraud. Based on the PRIDIT model, our PRIDIT score results show no significant difference in fraud patterns between the Takaful (Islamic) and commercial healthcare insurance markets. However, the PRIDIT weighting results indicate that the type of insurance (Takaful or Commercial insurance) can be relatively suspicious. Finally, we divide the data for fraud detection by visit type (outpatient versus inpatient) and find that the degree of suspicion is lower in the inpatient claims group than in the outpatient claims group. These results are consistent with the results of analyses in the US that have shown a lower rate of fraud activities in inpatient claims (e.g., Parente, et al., 2012).

Most of the current analysis focuses on understanding the relationships between different predictor variables and PRIDIT scores. Although we identify the most suspicious claims in our dataset, in future studies, we wish to relate these claims more directly to the cost of processing claims, settlement time, and claim disposition: a further analysis is required to determine the reasons for the rejection as well as to conduct a more granular analysis of the specific types of services that result in rejected claims. In addition, we plan to divide the analysis by ICD9 code and HCP type (e.g., doctor, emergency center, lab) to further refine our results. Again, we plan to conduct this additional analysis, which is beyond the scope of this paper, in future studies.

Our results indicate that evaluating and understanding the extent and potential drivers of fraud provide public policy implications for all parties in the insurance industry: policymakers (insurers), PHs and regulators. For example, regulators can set and/or update current regulations to better control fraudulent actions by PHs and HCPs. Policymakers can develop better strategies to investigate claims fairly and settle them efficiently. Furthermore, the results of our study and other healthcare fraud studies can help policymakers quantify the healthcare fraud rate and charge more accurate premiums (Ai et al., 2018). As a result, PHs will benefit from a more efficient healthcare insurance market by being able to pay a fairer premium.

Figure 3.1. Distribution of Claims Size over Healthcare Provider (HCP) Types



HCPs are (1) Pharmacy, (2) Doctor, (3) Emergency Center, (4) Inpatient Facility, (5) Lab, (6) Radiology Center, and (7) Medical Service Center

Figure 3.2. Distribution of Claim Size over Medical Coverage Types

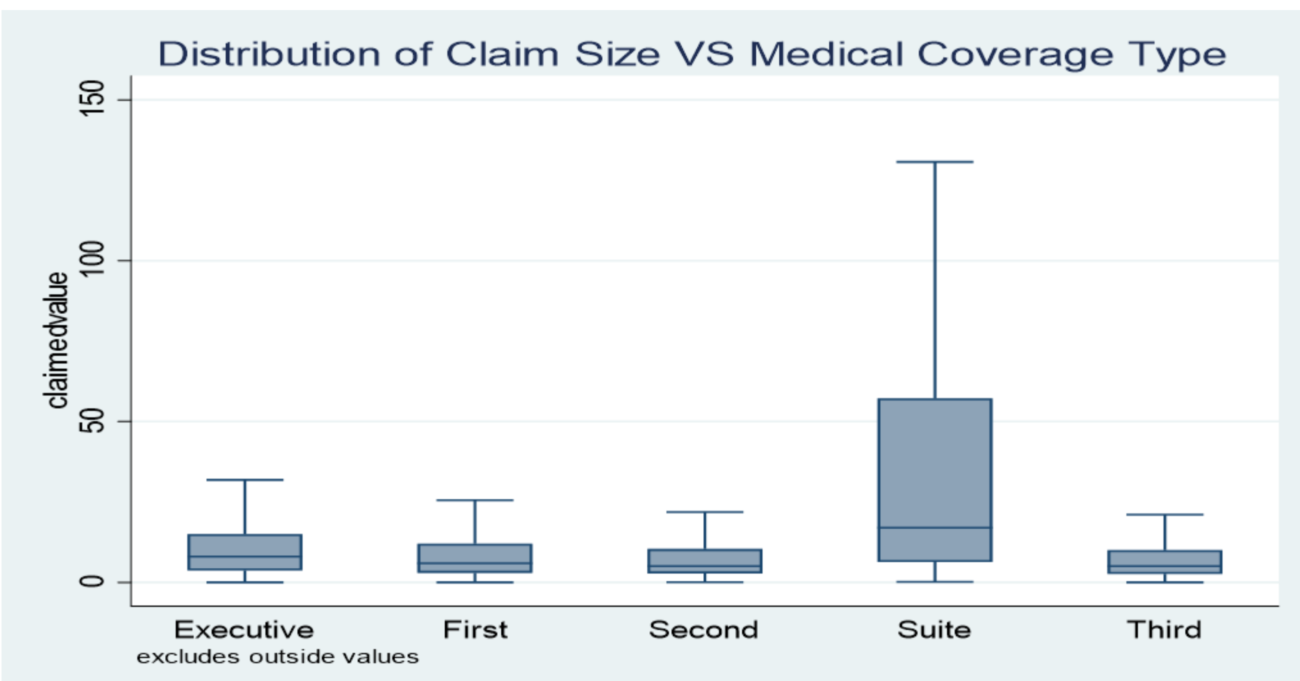
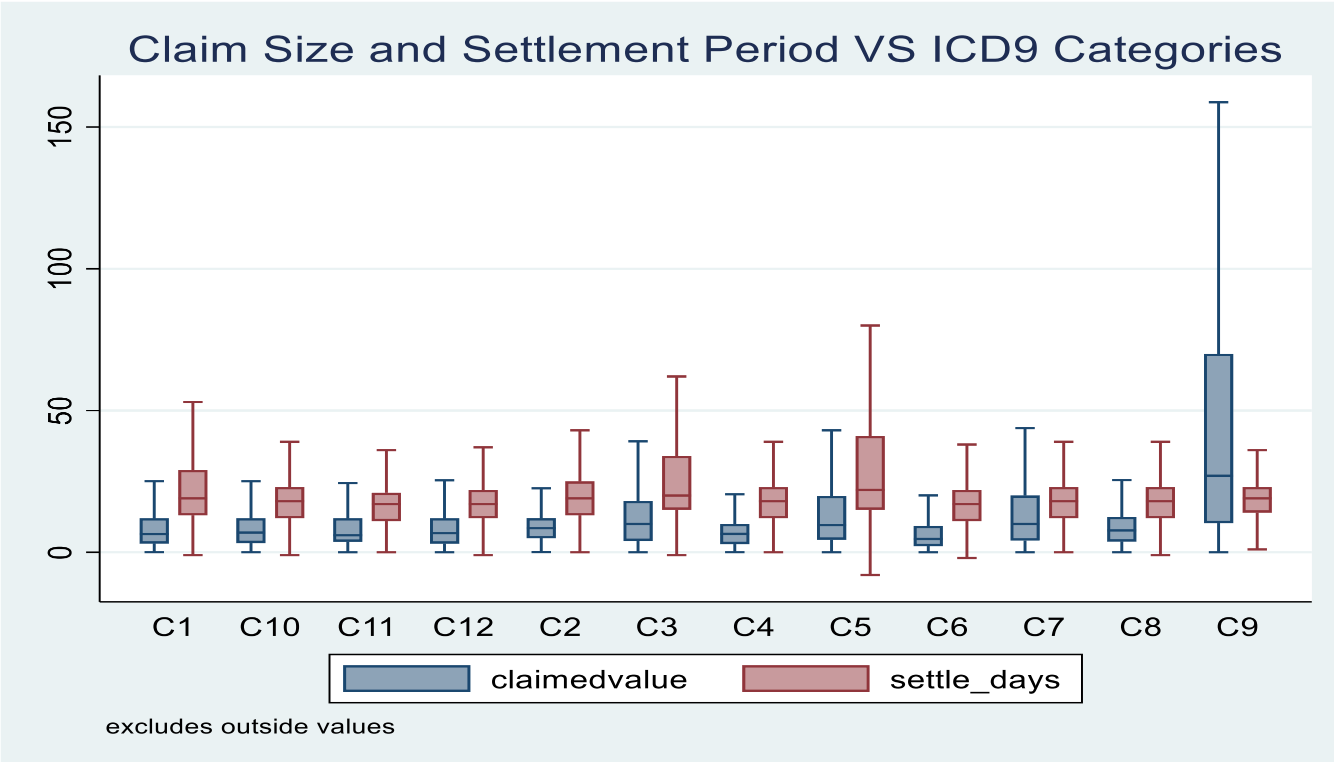


Figure 3.3. Distribution of Claim Size and Settlement Period across ICD9 Categories



C1: codes 001–139: infectious and parasitic diseases; **C2**: codes 140–239: neoplasms and codes 280–289: diseases of the blood and blood-forming organs; **C3**: codes 240–279: endocrine, nutritional and metabolic diseases and immunity disorders; **C4**: codes 290–319: mental disorders, codes 320–359: diseases of the nervous system and codes 360–389: diseases of the sense organs; **C5**: codes 390–459: diseases of the circulatory system; **C6**: codes 460–519: diseases of the respiratory system; **C7**: codes 520–579: diseases of the digestive system; **C8**: codes 580–629: diseases of the genitourinary system; **C9**: codes 630–679: complications of pregnancy, childbirth, and the puerperium, codes 740–759: congenital anomalies and codes 760–779: certain conditions originating in the perinatal period; **C10**: codes 680–709: diseases of the skin and subcutaneous tissue and codes 710–739: diseases of the musculoskeletal system and connective tissue; **C11**: codes 780–799: symptoms, signs, and ill-defined conditions; **C12**: codes 800–999: injury and poisoning and of E and V codes: external causes of injury and supplemental classification

.

References

Abdul Rahman, Z. (2009). Takaful: Potential demand and growth. *Islamic Economics*, 22(1).

Ai, J., Brockett, P. L., Golden, L. L., & Guillén, M. (2013). A robust unsupervised method for fraud rate estimation. *Journal of Risk and Insurance*, *80*(1), 121-143.

Ai, J., Lieberthal,R., Smith, S.,and Wojciechowski.R. (2018). Examining Predictive Modeling–Based Approaches to Characterizing Health Care Fraud. *Society of Actuaries*.

Ai, J., Brockett, P. L., & Golden, L. L. (2009). Assessing consumer fraud risk in insurance claims: An unsupervised learning technique using discrete and continuous predictor variables. *North American Actuarial Journal*, 13(4), 438-458.

Artís, M., Ayuso, M., & Guillen, M. (1999). Modelling different types of automobile insurance fraud behaviour in the Spanish market. *Insurance: Mathematics and Economics*, 24(1-2), 67-81.

Artís, M., Ayuso, M., & Guillén, M. (2002). Detection of automobile insurance fraud with discrete choice models and misclassified claims. *Journal of Risk and Insurance,* 69(3), 325-340.

American Medical Association (2010), AMA 2010 National health insurance report card, *American Medical Association Annual Meeting*, Chicago, IL., June 14. http://www.pnhp.org/news/2010/june/ama-health-insurer-report-card

Berwick, D. M., & Hackbarth, A. D. (2012). Eliminating waste in US health care. *Jama*, 307(14), 1513-1516.

Behrman, J. R., Mitchell, O. S., Soo, C. K., & Bravo, D. (2012). How financial literacy affects household wealth accumulation. *American Economic Review*, 102(3), 300-304.

Bhaskaran, S. R., & Puelz, R. (2009). Insurance Claim Operations: The Role of Economic Incentives. Available at SSRN 1522164.

Butler, S., & Francis, P. (2010). Cutting the cost of Insurance Claims: taking Control of the process, *Booz & Company*.

Brockett, P. L., Derrig, R. A., Golden, L. L., Levine, A., & Alpert, M. (2002). Fraud classification using principal component analysis of RIDITs, *Journal of Risk and Insurance*, 69(3), 341-371.

Childs, J. (2012). Gender differences in lying, *Economics Letters*, 114: 147-149.

Chui, A. C., & Kwok, C. C. (2008). National culture and life insurance consumption, *Journal of International Business Studies*, 39(1), 88-101.

Cremeans, K., Marcum, S., Followay, C., Oldaker, J., & Coustasse, A. (2019). Implications of upcoding on Medicare.Crocker, K. J., & Tennyson, S. (2002). Insurance fraud and optimal claims settlement strategies. *The Journal of Law and Economics*, 45(2), 469-507.

Cummins, J. D., & Tennyson, S. (1996). Moral hazard in insurance claiming: evidence from automobile insurance. *Journal of Risk and Uncertainty*, 12(1), 29-50.

Dreber, A. and M. Johannesson (2008). Gender differences in deception, *Economics Letters*, 99: 197-199.

Derrig, R. A. (2002). Insurance fraud. *Journal of Risk and Insurance*, 69(3): 271-287.

Derrig, R. A., & Krauss, L. K. (1994). First steps to fight workers' compensation fraud. *Journal of Insurance Regulation*, 12(3): 390-415.

Dionne, G., Giuliano, F., & Picard, P. (2009). Optimal auditing with scoring: Theory and application to insurance fraud. *Management Science*, 55(1), 58-70.

Fang, H., & Gong, Q. (2017). Detecting potential overbilling in Medicare reimbursement via hours worked. *American Economic Review*, 107(2), 562-91.

Friedman, Daniel. Morals and markets: An evolutionary account of the modern world. Springer, 2008.

Farbmacher, H., Löw, L., & Spindler, M. (2019). An Explainable Attention Network for Fraud Detection in Claims Management.

Friesen, L. and L.Gangadharan.(2012). Individual level evidence of dishonesty and the gender effect. *Economics Letters*, 117: 624–626.

Folland, S., A. C. Goodman, and M. Stano. (2016). The Economics of Health and Health Care. Pearson International Edition. New York: Routledge.

Geruso, M., & Layton, T. (2015). Upcoding: Evidence from Medicare on squishy risk adjustment (No. w21222). *National Bureau of Economic Research*.

Goates, S. (2010). Unintended Consequences: Medicare's Impact on the Diagnosis of Non-Enrollees.

Golden, L. L., Brockett, P. L., Guillén, M., & Manika, D. (2019). A PRIDIT Unsupervised Classification with Asymmetric Valuation of Variable Discriminatory Worth. *Multivariate Behavioral Research*, 1-19.

Guiso, L.,Sapienza,P.,Zingales,L.(2003). People's opium? Religion and economic attitudes. *Journal of Monetary Economics*, 50, 225–282.

Hilary, G. and Hui, K. W. (2009). Does religion matter in corporate decision making in America?, *Journal of Financial Economics*, 93, 455-473.

Iannaccone, L. R. (1998). Introduction to the Economics of Religion. *Journal of economic literature*, 36(3), 1465-1495.

Iyer, S. (2016). The new economics of religion. *Journal of Economic Literature*, 54(2), 395-441.

Insurance Research Council (1996). Fraud and Build-up in Auto Insurance Claims; Pushing the Limits of the Auto Insurance System. IRC.

Islamic Financial Services Industry Report. (2015). Islamic Financial Services Board, Malaysia [http:// www.ifsb.org]

Johnson, L. W. (2016). Federal Health Care Fraud Statute Sentencing in Georgia and Florida, 2011-2012.

Johnson, M. E., & Nagarur, N. (2016). Multi-stage methodology to detect health insurance claim fraud. *Health care management science*, 19(3), 249-260.

Joudaki, H., Rashidian, A., Minaei-Bidgoli, B., Mahmoodi, M., Geraili, B., Nasiri, M., & Arab, M. (2016). Improving fraud and abuse detection in general physician claims: a data mining study. *International journal of health policy and management*, 5(3), 165.

Kang, H., Hong, J., Lee, K., & Kim, S. (2010). The effects of the fraud and abuse enforcement program under the National Health Insurance Program in Korea. *Health policy*, 95(1), 41-49.

Kwon, W. J. (2007). Islamic principle and Takaful insurance: re-evaluation. *Journal of Insurance Regulation*, 26(1), 53.

Lin, K. C., & Yeh, C. L. (2012). Use of Data Mining Techniques to Detect Medical Fraud in Health Insurance. *International Journal of Engineering and Technology Innovation*, 2(2): 126-137.

Liou, F. M., Tang, Y. C., & Chen, J. Y. (2008). Detecting hospital fraud and claim abuse through diabetic outpatient services. *Health care management science*, 11(4), 353-358.

Major, J. A., & Riedinger, D. R. (2002). EFD: A hybrid knowledge/statistical-based system for the detection of fraud. *Journal of Risk and Insurance*, 69(3), 309-324.

Mesa, F. R., Raineri, A., Maturana, S., & Kaempffer, A. M. (2009). Fraud in the health systems of Chile: a detection model. Revista panamericana de salud publica= Pan American journal of public health, 25(1), 56-61.

Norris, P., & Inglehart, R. (2011). Sacred and secular: Religion and politics worldwide. Cambridge University Press.

Ortega, P. A., Figueroa, C. J., & Ruz, G. A. (2006). A Medical Claim Fraud/Abuse Detection System based on Data Mining: A Case Study in Chile. *DMIN*, 6: 26-29.

Outreville, J (2018). Culture and Life Insurance Ownership: Is It an Issue?. *Journal of Insurance Issues*, 41, 168-192

Park, S. C., and J. Lemaire (2012). The Impact of Culture on the Demand for Non- Life Insurance. *Astin Bulletin*, 42(2): 501–527

Parente, S. T., Schulte, B., Jost, A., Sullivan, T., & Klindworth, A. (2012). Assessment of predictive modeling for identifying fraud within the Medicare program. *Health Manag Policy Innov*, 1(2), 8-37.

Pande, V., & Maas, W. (2013). Physician medicare fraud: characteristics and consequences. *International Journal of Pharmaceutical and Healthcare Marketing*, 7(1), 8-33.

Picard, P. (1996). Auditing claims in the insurance market with fraud: The credibility issue. *Journal of Public Economics*, 63: 27–56.

Picard, P. (2013). Economic analysis of insurance fraud. *In Handbook of Insurance*, 315-316. Springer New York.

Shin, H., Park, H., Lee, J., & Jhee, W. C. (2012). A scoring model to detect abusive billing patterns in health insurance claims. *Expert Systems with Applications*, 39(8), 7441-7450.

Stavrova, O., & Siegers, P. (2014). Religious prosociality and morality across cultures: How social enforcement of religion shapes the effects of personal religiosity on prosocial and moral attitudes and behaviors. *Personality and Social Psychology Bulletin*, 40(3), 315-333.

Sharawy, H. M. (2000). Understanding the Islamic prohibition of interest: A guide to aid economic cooperation between the Islamic and Western worlds. Ga. J. Int'l & Comp. L., 29, 153.

Tennyson, S., & Salsas-Forn, P. (2002). Claims auditing in automobile insurance: fraud detection and deterrence objectives. *Journal of Risk and Insurance*, 69(3): 289-308.

Tennyson, S. (2008). Moral, Social, and Economic Dimensions of Insurance Claims Fraud. *Social Research*, 75(4): 1181-1204.

Thornton, D., Mueller, R. M., Schoutsen, P., & Van Hillegersberg, J. (2013). Predicting healthcare fraud in medicaid: a multidimensional data model and analysis techniques for fraud detection. Procedia technology, 9, 1252-1264.

Trinh, T., X. Nguyen, and P. Sgro (2016). Determinants of Non-Life Expenditure in Developed and Developing Countries: An Empirical Investigation. *Applied Economics* 48(58): 5639–5653.

Van Capelleveen, G., Poel, M., Mueller, R. M., Thornton, D., & van Hillegersberg, J. (2016). Outlier detection in healthcare fraud: A case study in the Medicaid dental domain. *International journal of accounting information systems*, 21, 18-31.

Weber, M. (2013). The Protestant ethic and the spirit of capitalism. Routledge.

Weisberg, H.I. and Derrig, R.A. (1991). Fraud and Automobile Insurance: A Report on the Baseline Study of Bodily Injury Claims in Massachusetts. *Journal of Insurance Regulation*, 9(4): 497–541.

Weisberg, H.I. and Derrig, R.A. (1992). Massachusetts Automobile Bodily Injury Tort Reform. *Journal of Insurance Regulation,* 10(3): 384–440.

Wojtusiak, J., Ngufor, C., Shiver, J., & Ewald, R. (2011). Rule-Based Prediction of Medical Claims' Payments: A Method and Initial Application to Medicaid Data. In Machine Learning and Applications and Workshops (ICMLA),10th International Conference, 2: 162-167.

Yang, W. S., & Hwang, S. Y. (2006). A process-mining framework for the detection of healthcare fraud and abuse. *Expert Systems with Applications,* 31(1): 56-68.

Appendix A: Calculating PRIDIT Scores, RIDIT Scores, and Their PRIDIT Weight

Table A1. Example of PRIDIT Scores and Their PRIDIT Weight for a Sample of Variables

Column Number	(1)	(2)	(3)= ∑(1)*(2) across All Variables
Variable	**PRIDIT Weight**	**RIDIT Score**	**PRIDIT Score**
Gender: (Female)	0.50	-0.75	-0.37 (Suspicious)
Dependents: (HOF)	0.40	0.65	0.26 (Nonsuspicious)
Marital status: (Single)	-0.45	0.95	-0.42 (Suspicious)
Age	0.85	0.40	0.34 (Nonsuspicious)
Overall PRIDIT Score			-0.20 (Suspicious)[80]

Table A2. Example of RIDIT Calculation for Gender

Column Number	(1)	(2)	(3)	(4)	(5)	(6)= (4)-(5)
Gender	**Category**	**Observations**	**Proportion**	**Proportion below Category**	**Proportion above Category**	**RIDIT Score**
Male	1	750	0.75	0	0.25	-0.25
Female	2	250	0.25	0.75	0	0.75

[80] The overall PRIDIT score, as explained previously, can also be positive (non-suspicious), depending on the sign/size of an individual's variable scores.

4. Chapter 4: To Leave or to Stay? Understanding the Effect of Switching between Medical Plans on Healthcare Insurance Adverse Selection

4.1 Introduction and Background

Insurance economists have widely documented that asymmetric information is present in many types of conventional insurance (Cohen and Siegelman, 2010).[81] Cohen and Siegelman (2010) note that insurance economists are on solid ground, as asymmetric information is present in most insurance markets. However, the authors call on insurance economists to change the direction of future research towards the challenges in separating the two sources of asymmetric information (i.e., separating adverse selection from moral hazard) and the drivers and contributors of asymmetric information.[82] In this paper, we answer that call by empirically investigating how switching between medical plans (up or down) interacts with adverse selection in an endogenous pricing setting in Takaful (Islamic) healthcare insurance, as religion and culture are crucial drivers of Takaful products. We also further extend the switching literature by considering how *adding* to a current policy (e.g., adding dental or vision coverage) affects variations in medical utilization asymmetry before and after *adding*. In this paper, we are interested in the effect of adverse selection on the intensive margin (i.e., does switching within Takaful schemes exacerbate adverse selection?).

Switching between medical plans may influence the extent of adverse selection and can also be a good predictor of future utilization (e.g., Altman et al. 1998, Cutler et al. 2010, Cabral 2016, and Cardon 2018). For instance, Tchernis et al. (2006) examine the switching behavior of a policyholder (PH) and healthcare utilization and find that those who switch to more generous medical plans tend to have higher utilization after switching than the nonswitchers who remain in the less generous plan (and vice versa).[83] The authors argue that the burden of adverse selection increases due to the switching costs associated with transferring from a less to a more generous medical plan.[84] Handel (2013) examines how an individual's medical plan choices affect adverse selection. He argues that reducing inertia increases the overall switching propensities of opportunistic switchers at the expense of nonswitchers, causing a reduction in

[81] Rowell, Nghiem, and Connelly, L. B. (2017) state that asymmetric information in insurance markets comprises two main components: (1) hidden information (adverse selection), which may be viewed as a pre-contractual form of information asymmetry) and (2) hidden action (moral hazard), which may be viewed as a post-contractual source of information asymmetry. For the purposes of this paper, we focus on adverse selection.
[82] Cohen and Siegelman (2010, p. 71) state, "The disentanglement of adverse selection and moral hazard is probably the most significant and difficult challenge that empirical work on adverse selection in insurance markets faces."
[83] Altman et al. (1998) reveal that adverse selection is influenced mainly by the current PHs' switching behavior, as the PHs' switching from generous medical plans to basic (economic) plans results in approximately 30-36 percent less spending (equivalent to $808 USD) by the switching PHs than that by those who choose to retain the most generous medical plans.
[84] Tchernis et al. (2006) further argue that switchers may postpone projected medical utilization/spending until after switching.

the nonswitchers' welfare (and exacerbating adverse selection). Cardon (2018) argues that switching between plans causes inefficiency and inequity, as some medical plans serve certain PHs at the expense of others in the form of a higher price. For instance, switching costs is considered a source of demand inelasticity in insurance markets, making them less competitive. Thus, it is critical to demonstrate a deep understanding of the extent and nature of such inefficiencies and inequities to implement mechanisms to reduce the deadweight "welfare loss" cost.

To date, how ADDING influences adverse selection in an existing medical policy has not been well documented and has never been explored across employer-sponsored (group) health plans, especially in the Takaful group health insurance market. For example, the literature has proposed only the following three factors that affect an individual's switching choice: (1) direct costs (e.g., premiums paid, out-of-pocket costs, and employer subsidies of perceived premiums), (2) the regulation of insurance markets (Polyakova 2016), and (3) expected future spending/utilization (health conditions) (Cutler et al. 2010). Furthermore, Cardon (2018) states that the PH's switching behavior can also be influenced by the indirect costs of switching (e.g., transaction costs, search/time costs, psychological costs, and choice compatibility costs), plan characteristics, and the PH's loss aversion.[85]

To examine the dynamic relationship between switching across plans and adverse selection, we utilize a unique insurance dataset from a Takaful (Islamic) health insurance company that contains switching/adding behavior (i.e., switching types and additions to policies), dates (i.e., switching dates and addition dates), premium, religion, and nationality for every individual PH. We first replicate the studies of Cardon (2018) and Cutler et al. (2010) on selected stories revealing an understanding of switching across medical plans to examine the relationship between switching across plans and adverse selection. As opposed to the dataset employed in Cutler et al. (2010), our dataset contains all additional subpolicies added by PHs to their insurance cart (e.g., optical or dental coverage that is not included in a basic policy). This enables us to evaluate the extent to which adverse selection changes when a current PH adds to a policy versus when a PH switches to another medical plan. The current literature has yet to examine how the PH's adding behavior influences adverse selection, especially in the Takaful group health insurance market. Similar to Cardon (2018), our work controls for the presence of moral hazard

[85] For more information about the rationale for switching between plans, see Cardon (2018), Handel (2013), Cutler et al. (2010), Tchernis et al. (2006), Cutler and Zeckhauser (2000), and Altman et al. (1998).

by construction because we compare medical utilization between switchers and nonswitchers within the same medical plans.[86]

We provide important evidence that adverse selection is present in the market and that switching costs implies asymmetry of medical utilization among switchers, nonswitchers, and those who add to their policy. Specifically, we show that the post-medical utilization of switchers, especially those who ADDED to their current medical policies, is economically larger than that of nonswitchers who remain in the HIGH or LOW medical plans. For instance, our results reveal that those who ADDED to a medical policy spent UAE dirham 163 (50 percent) more than those who remained in the LOW plan and 123 (31 percent) more than those who stayed in the HIGH plan. Furthermore, the switchers from LOW to HIGH plans (switch UP) spent UAE dirham 263 (68 percent) more than those who stayed in a LOW plan and 223 (49 percent) more than those who stayed in a HIGH plan. Our results show that the effects are even larger than those found by Cardon (2018). For instance, he shows that those who switched up spent 23 percent more than those who stayed HIGH. Furthermore, those who switched down spent 4.7 percent less than those who stayed LOW.

In this paper, we contribute to the existing insurance and health economics literature that examines adverse selection in healthcare insurance markets (e.g., Polyakova, 2016; and Cohen and Siegelman 2010)[87] and the determinants of the PH's switching behavior (e.g., Cardon, 2018; Handel, 2013; Cutler et al., 2010; Tchernis et al., 2006; Cutler and Zeckhauser, 2000; and Altman et al., 1998). Specifically, this paper fills a gap in the literature by documenting the effect of adding to a current medical policy on adverse selection and PH medical utilization, as adverse selection is indicated by additions. Adding is important relative to switching, as our data show that 23 percent of PHs add to their current medical policy, 12 percent switch down, and 2 percent switch up (see Table 2).[88] The current literature has yet to examine the dynamic relationship between the PH's adding behavior and their medical utilization. Furthermore, there is an apparent need for more research on the factors and circumstances that influence the existence of adverse selection (i.e., the dynamic relationship between switching costs and adverse selection). Some researchers have examined such interactions (switching costs and adverse selection)

[86] Chipper and Salanie (2000) report that a positive correlation between claim performance variables (i.e., size and number of claims) and insurance coverage indicates the possibility of adverse selection and asymmetric information. Their tests do not disentangle adverse selection and moral hazard. However, Cardon (2018) reports that comparing medical spending between switchers and nonswitchers within the same medical plan implicitly controls for the presence of moral hazard.

[87] For more information about asymmetry and adverse selection in insurance markets, see Einav, Finkelstein, and Levin (2010), and Cohen and Siegelman (2010).

[88] Those who switched up paid on average UAE dirham 502 more for premiums than those who added. Furthermore, those who added paid on average UAE dirham 2,710 less for premiums than those who stayed high. This finding suggests that those who add are better off than those who stay high, considering their medical utilization.

within public and social healthcare insurance markets (e.g., Medicare).[89] To the best of our knowledge, this paper is the first to examine switching costs and evaluate the potential interaction between switching behavior, adding, and adverse selection in a private Takaful health insurance market, where employees and their dependents are covered by employer-sponsored health insurance.[90]

The remainder of this paper proceeds as follows. The next section provides a detailed theoretical background and develops our hypotheses. The third section describes the data and methodology used to test our hypotheses. The fourth section presents our results, and the final section offers concluding remarks and suggestions for future research.

4.2 Theoretical Background and Development of the Hypotheses

How PHs behave is important to insurance companies because individuals have information about their level of risk that is unknown to insurers. According to theory, this information asymmetry can lead to adverse selection[91] (Acerola, 1970; Rothshild and Stiglitz, 1976).[92]

After the seminal contributions of Akerlof (1970) and Rothshild and Stiglitz (1976), researchers have empirically examined the presence of adverse selection by testing the relationship between insurance demand and insurance utilization (measured by claim behavior). Significant research in this context provides evidence of adverse selection in healthcare insurance markets (e.g., Panhans 2019, Cutler and Zeckhauser 2000, Simon 2005, Eling, Jia, and Yao 2015). Cohen and Siegelman (2010) provide a detailed literature review of empirical studies that focus on conventional insurance markets across different business lines. They note that empirical research in this area has generated mixed results across all commercial insurance lines, and most insurance economists are on solid ground in believing that asymmetric information and/or selection exist in most insurance markets.[93] As such, we posit the following hypotheses.

---

[89] Polyakova (2016) utilizes the variation of prices and regulation in Medicare Part D and empirically examines the interactions between switching costs, regulation, and adverse selection.

[90] Tchernis et al. (2006) report that approximately 60 percent of individuals in the USA are insured via employer-sponsored health insurance and that approximately 50 percent of that group can choose from medical plans that vary in their level of generosity.

[91] Cummins, Smith, Vance and VanDerhei (1983) define adverse selection as "… the tendency of high-risks to purchase insurance or to purchase more insurance coverage than do low-risks."

[92] For instance, the theoretical work of Rothschild and Stiglitz (1976) provides evidence that profit-maximizing policymakers are incapable of distinguishing high risks from low risks in a competitive market. This would make low-risk policies worse off due to the presence of high-risk individuals, as it would motivate the low-risk individuals to buy less insurance than they would in a market where asymmetric information is not present.

[93] For more information about the empirical work on adverse selection, see Cohen and Siegelman (2010) and the first essay of this dissertation (Section 1.4).

H1. Adverse selection is present in the healthcare insurance market, where PHs switch between medical plans and/or add to a current medical policy.

H1A. Adverse selection is significantly more prevalent among PHs who add and/or switch up than those who switch down, stay low and/or stay high.

Next, we consider how adding to a current medical policy influences the presence of adverse selection. The literature has made no assumptions about the prevalence of adverse selection before and after adding to a current medical policy. Insurance common sense suggests that compared to the extent of adverse selection among policyholders who never add to their policies, , t, the extent of adverse selection among policyholders who add to their current policies will increase. Furthermore, the prevalence of adverse selection before and after switching between medical plans (e.g., switching UP from an economical medical plan to a VIP plan) requires further examination (see Cardon, 2018). Our hypotheses are developed from empirical research that examines the drivers and contributors of asymmetric information (e.g., Elin, Jia, and Yao 2015) as well as the literature on switching costs in insurance markets (e.g., Handel 2013, Cutler, Lincoln, and Zeckhauser 2010, and Cardon 2018).

Economists have already examined the effect of medical plan switching on adverse selection and have argued that switching between medical plans predicts future utilization (e.g., Cardon 2018, Handel 2013, Cutler et al. 2010, Tchernis et al. 2006, Cardon and Hendel 2001, Robinson et al. 1993). Specifically, Tchernis et al. (2006) report significant differences in medical expenditures before and after plan switching. They find that medical utilization expenditures for those who switch down to a basic economic plan (i.e., a less generous plans) are significantly lower than the expenditures for those who never switch and stay in a more generous plan (and vice versa). The results are consistent with Robinson et al. (1993), who report that individuals anticipate maternity care before switching to more generous medical plans. This finding suggests that PHs tend to delay medical utilization until after switching to higher medical plans.

Cutler et al. (2010) explain the stories and economic rationale behind moving between medical plans. They argue that adverse selection is not the only motive for switching between medical plans[94] and provide evidence that switching costs implies asymmetry of medical utilization between switchers and nonswitchers, as switching to a more generous medical plan is riskier than staying in a highly generous

[94] Adverse retention and inertia (aging in place) are also potential cost-related variables related to switching. However, given the limitations of our dataset, we focus on adverse selection. For more information about adverse retention, see Cutler et al. (2010).

plan. However, switching to low-coverage plans is less risky than staying in comparable plans. The asymmetry of medical utilization is a result not only of adverse selection but also of the interaction between asymmetric information and switching costs.

Given the existing literature on how switching costs influences adverse selection, we expect to find that the prevalence of adverse selection is larger among switchers than among nonswitchers. Thus, we hypothesize the following.

H2A. The extent of adverse selection is lower among nonswitchers than among switchers.

H2B. Post-switch healthcare expenditures for switchers who SWITCH UP are larger than the expenditures for nonswitchers who remain in HIGH or LOW medical plans.

H2B. Post-switch healthcare expenditures for switchers who SWITCH DOWN are lower than the expenditures for nonswitchers who remain in HIGH or LOW medical plans.

Next, we consider the variation in medical utilization asymmetry before and after adding to a current policy. Adding to a current medical policy may suggest that healthcare services covered by these plans complement the existing medical policy. Thus, one would expect that an individual's utilization would increase after adding to an existing medical policy.

We argue that subsequent medical utilization is greater for those who ADD than for those who never ADD and/or for those who remain in HIGH or LOW medical plans. Adding to a current medical policy and/or switching UP to a more generous medical plan may be riskier than remaining in a HIGH or LOW medical plan. We argue that the presence of adverse selection and/or asymmetric information is more pronounced among switchers than among nonswitchers and is even more pronounced among those who ADD/SWITCH UP than among those who switch DOWN or stay in LOW/HIGH medical plans. Thus, we formally postulate the following.

H3. Post-switch healthcare expenditures for those who ADD to a current medical policy are larger than the expenditures for nonswitchers who remain in HIGH or LOW medical plans.

4.3 Data and Methodology

To test our hypotheses, we use a proprietary dataset that contains all PHs and individual medical claims from private insurance companies operating in the United Arab Emirates (UAE)[95] from 2014 to 2015.

[95] The company is representative of the UAE insurance market, as it has a large market share among Takaful and commercial companies.

Our data are useful for investigating switching costs and adverse selection[96] for the following three reasons.

- *First*, this dataset contains information about PHs from 141 different countries, but they all reside in the UAE. Specifically, the insurance company provided us with information about individual PHs' insurance coverage and medical utilization (i.e., claim performance variables, such as total claims filed and the size of each claim). These detailed variables enable us to examine whether adverse selection is present in healthcare insurance markets. Chipper and Salanie (2000) report that a positive correlation between claim performance variables and insurance coverage suggests that high-risk individuals tend to buy insurance with more generous coverage. This positive correlation indicates the possibility of adverse selection and/or asymmetric information. A nonsignificant correlation between risk (i.e., claim performance variables) and insurance coverage may suggest that no adverse selection and/or asymmetric information exist.
- *Second*, this dataset contains the PHs' switching type (i.e., switch up, switch down, or add to a current policy), switching date, characteristics of each PH's insurance policy (i.e., policy commencement and termination dates), and medical claim characteristics for each insurance claim made (i.e., filing date). With these various dates, we can create a panel of insurance claim behavior for enrolled PHs and then examine the effect of switching behavior on adverse selection.
- *Third*, this dataset contains the PHs' medical visit characteristics (i.e., inpatient versus outpatient, disease type (ICD9 codes)) and PH characteristics (i.e., age, gender, education, dependents, and marital status). With these various unique variables, this study can test our hypotheses by simultaneously controlling for the effects of demographic and medical visit characteristics.

Table 4.1 lists the variables utilized in this paper alongside brief descriptions. To test the possibility of adverse selection, we examine the association between medical expenditure (i.e., risk) and insurance coverage. We use the following two measures of claim performance as a measure of risk and medical utilization: (1) the amount claimed by the PH and (2) the total number of claims filed by the PH. As a proxy for insurance coverage, we take the natural logarithm of the annual maximum level of insurance coverage. We include the following eight control variables: (1) gender, (2) age, (3) relationship to the principal PH (i.e., spouse or child), (4) marital status, (5) educational background, (6) policy length (in

[96] Insurance economists have reported that disentangling adverse selection from moral hazard is an empirical challenge due to data limitations (e.g., Cohen and Siegelman, 2010). Thus, the focus of our study is adverse selection. However, Cardon (2018) reports that comparing medical spending between switchers and nonswitchers within the same medical plans implicitly controls for the presence of moral hazard.

months) for each beneficiary under an insurance policy, (7) visit type (inpatient versus outpatient), and (8) disease type (ICD9 code).

Table 4.1. Descriptions of Variables	
Variable	Description
<u>Dependent Variable (Medical Expenditure):</u>	
Medical Expenditure	Claim amount paid to beneficiary
Number of Claims	Total number of claims filed by the policyholder (PH)
<u>Key Variables of Interest</u>	
Coverage Amount	Natural logarithm of the maximum amount of insurance coverage
Switch Up	1 if a PH switches up to a generous plan (i.e., VIP plan) and 0 otherwise
Switch Down	1 if a PH switches down from a generous plan (i.e., VIP plan) to a basic plan (i.e., economy plan) and 0 otherwise
Add	1 if a PH adds to a current policy and 0 otherwise
Stay High	1 if a PH never switches and stays in a generous plan and 0 otherwise
Stay Low	1 if a PH never switches and stays in a basic plan (i.e., economy plan) and 0 otherwise
<u>Control Variables</u>	
Policy Length	Length of policy in months
Age	Age of beneficiary in years [?]
Male	1 if male and 0 otherwise
Married	1 if married and 0 otherwise
Principal	1 if the principal of the policy (head of family) and 0 otherwise
Spouse	1 if the spouse of the PH and 0 otherwise
Child Dependent	1 if the child is a dependent of the PH and 0 otherwise
College Degree	1 if the PH is educated (college degree) and 0 otherwise
Chronic	1 if the claim is made for a chronic condition and 0 otherwise
Outpatient	1 if the visit type is outpatient and 0 otherwise
Inpatient	1 if the visit type is inpatient and 0 otherwise
<u>Classification of Diseases (ICD9 Chapters)</u>	
ICD9_Cat1	1 for diseases of the respiratory system: ICD-9 codes 460–519 and 0 otherwise. *<u>ICD9_Cat1 is the admitted variable</u>*
ICD9_Cat2	1 for infectious and parasitic diseases: ICD-9 codes 001–139 and 0 otherwise
ICD9_Cat3	1 for dermatologic, musculoskeletal and rheumatologic disorders: ICD-9 codes 680–709, ICD-9 codes 710–739 and 0 otherwise
ICD9_Cat4	1 for symptoms, signs, and ill-defined conditions: ICD-9 codes 780–799 and 0 otherwise
ICD9_Cat5	1 for injury and poisoning diseases: ICD-9 codes 800–99, ICD-9 E and V codes and 0 otherwise
ICD9_Cat6	1 for neoplasms, hematology and oncology disorders: ICD-9 codes 140–239, ICD-9 codes 280–289 and 0 otherwise
ICD9_Cat7	1 for endocrine, nutritional and metabolic diseases and immunity disorders: ICD-9 codes 240–279 and 0 otherwise
ICD9_Cat8	1 for mental/psychiatric disorders, diseases of the nervous system, and sense organs: ICD-9 codes 290–319, ICD-9 codes 320–359, ICD-9 codes 360–389 and 0 otherwise
ICD9_Cat9	1 for diseases of the circulatory system: ICD-9 codes 390–459 and 0 otherwise
ICD9_Cat10	1 for diseases of the digestive system: ICD-9 codes 520–579 and 0 otherwise
ICD9_Cat11	1 for diseases of the genitourinary system: ICD-9 codes 580–629 and 0 otherwise
ICD9_Cat12	1 for complications of pregnancy, childbirth, and the puerperium and for congenital anomalies and certain conditions originating in the perinatal period: ICD-9 codes 630–679, ICD-9 codes 740–759, ICD-9 codes 760–779 and 0 otherwise

Table 4.2 reports the summary statistics of our dataset. The table reveals that the average medical expenditure (claimed amount) for all PHs is approximately UAE dirham 421, which is equal to approximately USD $126, the average number of claims filed is 24.6, and the average insurance coverage amount is UAE dirham 906,408 (USD $246,779). Regarding our key variables, Table 4.2 shows that 12 percent of PHs switched down, 2 percent switched up, and 23 percent added to their current medical policy. The PHs in our sample had access to five medical plan options. The first two are generous medical plans with a comprehensive medical network (i.e., the "VIP Plan" and the "Regional Plus Plan"). The last three are significantly less generous, with a restricted medical network (basic plan, economy plan, and regional plan). We group the two generous plans (HIGH) and compare them with the group of three low plans (LOW). Table 4.2 shows that 34 percent of the PHs stay in the high plans (Stay High), and 27 percent of PHs stay in the less generous plans (Stay Low).

Regarding PH characteristics, Table 4.2 shows that 67 percent of the PHs are male, 72 percent are married, 16 percent have children, the average age is 32 years, and 58 percent are college educated. Finally, the PHs' policy/medical visit characteristics are as follows: 91 percent of PHs file outpatient claims; the average policy length for all PHs is 20 months; 16 percent have chronic conditions; 23 percent suffer from diseases of the respiratory system; 17 percent suffer from dermatologic, musculoskeletal and rheumatologic disorders; and 12 percent suffer from diseases of the digestive system. Thus, it is important to use previous demographic and medical visit measures as controls.

Table 4.2. Summary Statistics

Column	1	2	3	4	5
	N	Mean	Std. Dev.	25th Percentile	75th Percentile
VARIABLE					
Medical Expenditure					
Medical Expenditure	556,201	421.27	1066.60	177.24	494.66
Number of Claims	556,201	22.58	22.71	7	31
Coverage Amount	556,201	870,018.6	700617	250,000	1,000,000
Switch Category					
Stay High	556,201	0.33	0.47	0	1
Stay Low	556,201	0.28	0.44	0	1
Add	556,201	0.24	0.42	0	0
Switch Down	556,201	0.13	0.32	0	0
Switch Up	556,201	0.02	0.15	0	0
College Degree	370,052	0.59	0.49	0	1
Age	556,201	32.33	15.66	26	42
Male	556,201	0.67	0.46	0	1
Married	556,201	0.72	0.44	0	1
Chromic	556,201	0.16	0.37	0	0
Principal	556,201	0.65	0.47	0	1
Spouse	556,201	0.18	0.38	0	0
Child	556,201	0.16	0.37	0	0
Outpatient	556,201	0.91	0.27	1	1
Inpatient	556,201	0.08	0.27	0	0
Policy Length	556,201	20.36	3.28	18	21
ICD9 Chapters					
ICD9_Cat1	556,201	0.23	0.42	0	0
ICD9_Cat2	556,201	0.03	0.18	0	0
ICD9_Cat3	556,201	0.17	0.37	0	0
ICD9_Cat4	556,201	0.11	0.31	0	0
ICD9_Cat5	556,201	0.06	0.24	0	0
ICD9_Cat6	556,201	0.01	0.10	0	0
ICD9_Cat7	556,201	0.04	0.20	0	0
ICD9_Cat8	556,201	0.06	0.24	0	0
ICD9_Cat9	556,201	0.04	0.21	0	0
ICD9_Cat10	556,201	0.12	0.32	0	0
ICD9_Cat11	556,201	0.05	0.23	0	0
ICD9_Cat12	556,201	0.03	0.17	0	0

To identify the extent of the effect of switching costs on adverse selection and to avoid the potential confounding effects of moral hazard, we compare the significant differences in medical expenditure before and after switching (down and up) and before and after adding to a current medical policy in the same plans. Following Cardon (2018), we argue that this approach "…compares spending by switching type within the same insurance plan [and] controls for the effect of moral hazard by construction." We then compare the differences in medical utilization between switchers and nonswitchers (e.g., those who remain in HIGH and LOW medical plans). Table 4.3 reports medical expenditure by switching category. We compare the significant differences in the average claim amount and number of claims filed between switchers (before and after switching) and nonswitchers. Columns (1) to (3) show the average medical expenditures and number of claims for HIGH medical plans (i.e., those who switched up, added to their medical policy, and stayed in highly generous plans, respectively). Columns (4) and (5) show the average medical expenditures for LOW medical plans (i.e., those who switched down and stayed in low economy plans, respectively).

Table 4.3 reveals that the medical expenses are generally quite different among switchers, especially between those who switched up (Column 1) and those who added to their medical policy (Column 3). For example, Column (1) shows that the post-switch medical expenditures for those who switched up averaged UAE dirham 801, compared with an average of UAE dirham 567 for their pre-switch medical expenditures, a difference of UAE dirham 234. The results are unchanged when we consider the average number of claims filed before and after switching up. The post-switch number of claims averaged 37, compared to the average of 32 claims filed before switching, a difference of 5 claims. When we examine pre- and post-addition medical expenditures, Column (3) shows consistent results, revealing that post-addition medical expenditures averaged UAE dirham 695, compared with an average of UAE dirham 446 before the addition, a difference of UAE dirham 249. Interestingly, the post-addition number of claims averaged 34, compared to 19 claims before the addition, a difference of 15 claims. The difference between these groups is statistically significant at the 1 percent level. This finding suggests that the prevalence of adverse selection is more pronounced among those who add to than those who switch up to more generous plans and that those who add to their medical policy may be more likely to engage in behavior representative of adverse selection.[97]

Table 4.3. Medical Expenditures by Switching Type

[97] Notably, the average premium for those who switched up was UAE dirham 3,592 and for those who added was UAE dirham 3,090 (a difference of UAE dirham 502). This suggests that considering their medical utilization, those who add are better off than those who switch up.

Table 4.3 reports medical expenditures by switching category. We compare the significant differences in the average claimed amount and number of claims filed between switchers (pre-/post-switch) and nonswitchers. Columns (1) to (3) show the average medical expenditures and number of claims for HIGH medical plans (i.e., those who switched up, added to their medical policy, and stayed in highly generous plans, respectively). Columns (4) and (5) show the average medical expenditures for LOW medical plans (i.e., those who switched down and stayed in low economy plans, respectively). The mean (t-test) differences are reported, and *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Column	(1)	(2)	(3)	(4)	(5)
	High (Generous Coverage)			Low (Economy Coverage)	
Mean	**Switch Up**	**Stay High**	**Add[98]**	**Switch Down**	**Stay Low**
a) Post-switch Medical Expenditure	801.22	511.91	659.40	415.17	251
b) Pre--switch Medical Expenditure	567.02	-	446.66	385.57	-
c) Post-switch Number of Claims	37	29	34	21	15
d) Pre-switch Number of Claims	32	-	19	20	-
Observations	13,411	199,672	183,523	87,182	171,344
Mean (t-test) differences (a) - (b)	234.2***	-	212.74***	29.6	-
Mean (t-test) differences (c) - (d)	5**	-	15***	1	-

The results in Column (4) show that there are no significant differences in either medical expenditures or the number of claims before and after switching down. These results suggest that PHs might have selected the wrong medical plan (e.g., overestimating their medical expenditures) and that they pay larger premiums (UAE dirham 5,800) for unneeded medical plans (stay HIGH). In such cases, switching down to less generous plans may be a rational (i.e., expected utility-maximizing) decision. Our data show that the average premium saved after switching down is UAE dirham 2,632 (5,800 minus 3,168, approximately USD $716).

Next, we consider the differences in medical utilization between switchers and nonswitchers (e.g., those who switched and added versus those who stayed HIGH or LOW and never switched). When we compare those who switched up (Column 1) to those who stayed LOW and never switched (Column 5), the results show that medical expenditure for switchers (UP) averaged UAE dirham 801, compared with UAE dirham 251 for those who remained in LOW plans (stay LOW), a difference of UAE dirham 551. Furthermore, the number of claims for switchers (UP) averaged 37, compared with 15 claims for nonswitchers (remaining in LOW plans), a difference of 22 claims. Interestingly, when we compare the results for those who stayed HIGH (Column 2), the medical expenditure for those remaining in HIGH

[98] Our dataset suggests that those who ADD belongs to the high coverage.

plans averaged UAE dirham 511, compared with UAE dirham 801 for those who switched up, a difference of UAE dirham 290. The difference in the number of claims for both groups is 8 claims (37 for switching to a high plan versus 29 for staying HIGH). This result suggests that adverse selection and/or asymmetric information is more pronounced among those who switch up than among those who stay HIGH. Hence, those who switch up are more likely to engage in behavior that reflects information asymmetry.

The results are qualitatively similar when we compare those who add and switch down (Columns 3 and 4, respectively) with those who stay LOW (Column 5). Post-addition expenditures were significantly higher than expenditures for those staying in LOW plans (UAE dirham 659 versus 251), and switchers and adders filed more claims (34 claims versus 15). On the other hand, switchers (from HIGH to DOWN) spent more than nonswitchers staying in LOW plans (415 versus 251). This result is not consistent with Cardon (2018), who reports that switchers (from HIGH to DOWN) spend less than nonswitchers remaining in LOW plans.

Overall, our results suggest that the presence of adverse selection and/or asymmetric information is more pronounced among switchers than among nonswitchers and is even more pronounced among those who add than among those who switch up (from LOW to HIGH). Furthermore, adverse selection and/or asymmetric information are more prevalent among those who switch up than among those who stay HIGH. The prevalence of adverse selection pre- and post-addition has not been previously explored by insurance economists. However, our results before and after switching up and down are consistent with Cardon (2018), who examines the effect of loss aversion on medical plan switching behavior. The author reports that medical expenditures are significantly larger for switchers (from LOW to HIGH) than for her comparison group pre- and post-switch. She states that "… this hints that the cost of switching up is higher than the cost of switching down." Given the differences among those groups, we encourage insurance economists to further explore the potential effect of switching behaviors on the extent of adverse selection, especially among those who add to their policy instead of switching up or down.

4.3.1 Methodology

To examine the relationship between switching costs and adverse selection, we employ an empirical strategy similar to that used in Cardon (2018) and estimate the relationship between switching costs and medical expenditures, controlling for individual observables. Cardon (2018) regresses total medical expenditures and the natural LOG of medical expenditures on switching categories and various demographic control variables. However, his dataset has a shortcoming in that it has no information about claiming behavior for PHs who add to their current policy and never switch up or down. As opposed to Cardon's (2018) dataset, our dataset contains all PH switching types (especially for those who add to a current policy).

Our model also allows us to test for the presence of adverse selection and is similar to that used in Finkelstein and Poterba (2004), who estimate the relationship between insurance coverage and claim behavior. More specifically, we estimate the following linear model for individual *i* on day *t*:

$$Expenditures_{it} = \alpha + \Omega lnCovAmt_i + \theta X_{it} + \beta R_i + \varepsilon_{it} \tag{4.1}$$

We define *Expenditures* as one of the two claim variables discussed in the previous Section 4.3 (total claimed amount and the total number of claims filed). Next, *lnCovAmt* is the logged yearly maximum insurance coverage level. We include *X* as a vector of individual-specific observable characteristics. Included in *X* are gender, age, marital status, policy length, relationship to the principal of the policy, education, chronic condition, visit type (inpatient or outpatient) and disease type (ICD9 chapters). The coefficient of interest for this regression is *β*. When considering the relationship between switching costs and adverse selection, we define *R* as one of the key switching variables described in the previous Section 4.3 (see Table 4.1).

The second coefficient of interest for this regression is **Ω**, which is used to test for the presence of adverse selection. According to Equation 4.1, if adverse selection is present, we expect that **Ω**>0. It is important to note that **Ω** has no causal interpretation due to the nature of adverse selection and asymmetric information. Chiappori and Salanie (2000) report that a positive correlation between claim performance variables (i.e., medical *expenditure* variables in our model) and insurance coverage suggests that high-risk individuals buy insurance with more generous coverage (e.g., switching up from a low medical plan to a more generous plan). This positive correlation indicates the possibility of adverse selection and asymmetric information. A nonsignificant correlation between risk and coverage suggests no adverse selection and asymmetric information are not present.

As an additional robustness check, we follow Finkelstein and McGarry (2006) and use the following two reduced-form models to examine the effect of switching costs on the presence of adverse selection:

$$Claim_{it} = \alpha + \theta X_{it} + \beta R_i + \varepsilon_{it} \tag{4.2}$$

$$lnCovAmt_{it} = \alpha + \theta X_{it} + \beta R_i + \gamma_{it} \tag{4.3}$$

According to these models, the positive correlation between the residuals ϵ and γ is a necessary condition for the presence of adverse selection or asymmetric information, which suggests that the relationship between coverage choice and medical expenditure (claim size and claim occurrence) is dependent after controlling for observables. A nonsignificant correlation between ϵ and γ suggests that there is no adverse selection or asymmetric information. Thus, we use Cardon's (2018) model as our core model. We then follow Finkelstein and McGarry (2006) and Chiappori and Salanie (2000) in using the reduced-form models as a robustness test.

3.4 Results

Table 4.4 presents the ordinary least squares (OLS) regression results of our core model. Columns 1 and 3 report the results for total expenditure and the number of claims, respectively. Columns 2 and 4 show the results for the natural log of total expenditure and number of claims, respectively. As discussed above, we define the following five categories of switching type: (1) switch up, (2) switch down, (3) ADD, (4) stay high, and (5) stay low. Focusing on the presence of adverse selection, we can see across all specifications (Columns 1 to 4) that the correlation between medical utilization variables and logged maximum insurance coverage is positive and significant at the 1 percent level. These results suggest that adverse selection is present in healthcare insurance markets.

To observe how switching between plans influences the extent of adverse selection, we need to focus on the results of asymmetric medical expenditures by switchers and nonswitchers. Column 1 shows that switchers from LOW to HIGH plans (switch UP) spent UAE dirham 263 more than those who stayed LOW and 223 more than those who chose to stay HIGH. Regarding switchers from HIGH to LOW (switch DOWN), Column 1 shows that they spent UAE dirham 101 more than those who stayed LOW. All of the results are significant at the 1 percent level. However, the switch down results are inconsistent with Cardon (2018), who reports that such switchers spend significantly less than those who remain LOW.

Table 4.4. The Effect of Switching Cost on Adverse Selection

The table shows the coefficients of interest for the ordinary least squares (OLS) regression results, estimating the relationship between our measures of medical expenditure, measures of switching costs, and logged maximum insurance coverage. The key variables of interest are as follows: (1) switch up is a binary variable with a value of one if the policyholder (PH) switched up to a more generous plan; (2) switch down is a binary variable with a value of one if the PH switched down to a basic (economy) plan; (3) addition is a binary variable with a value of one if the PH added to a current policy; (4) stay high is a binary variable with a value of one if the PH remained in a generous medical plan; and (5) stay low is a binary variable with a value of one if the PH remained in a basic (economy) medical plan (stay low is the base). We include age, gender, marital status, relationship to the principal, policy length, chronic condition, visit type, and disease type (ICD9 categories) as controls in all regressions. For the control variable descriptions, refer to Table 1. Robust standard errors are reported in parentheses, and *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Variable	(1) Total Exp.	(2) log(Total Exp.)	(3) Number of Claims	(4) log(Number of Claims)
Coverage	104.4***	0.256***	5.764***	0.311***
	(3.623)	(0.00146)	(0.0543)	(0.00231)
Stay High	39.92***	0.189***	0.614***	-0.0787***
	(5.426)	(0.00292)	(0.0984)	(0.00436)
Addition	163.5***	0.503***	1.085***	-0.0818***
	(4.580)	(0.00244)	(0.0765)	(0.00401)
Switch Down	101.1***	0.266***	-0.183*	-0.160***
	(10.17)	(0.00308)	(0.0976)	(0.00486)
Switch Up	263.3***	0.684***	6.191***	0.104***
	(11.29)	(0.00510)	(0.243)	(0.00822)
Inpatient	399.3***	0.381***	-1.902***	-0.0681***
	(28.28)	(0.00487)	(0.141)	(0.00576)
Chronic	74.93***	0.127***	3.000***	0.182***
	(10.96)	(0.00270)	(0.100)	(0.00375)
Age	6.535***	0.0108***	0.284***	0.0125***
	(0.127)	(9.36e-05)	(0.00368)	(0.000133)
Male	-6.348	-0.0486***	-5.154***	-0.199***
	(5.608)	(0.00242)	(0.0965)	(0.00358)
Married	-19.65***	-0.000272	4.396***	0.251***
	(4.024)	(0.00275)	(0.0688)	(0.00409)
Spouse	54.86***	0.114***	2.279***	0.140***
	(4.244)	(0.00298)	(0.136)	(0.00440)
Child	71.09***	0.0334***	5.461***	0.453***
	(7.178)	(0.00433)	(0.139)	(0.00636)
Policy Length	4.321***	0.00960***	0.928***	0.0368***
	(0.380)	(0.000249)	(0.00872)	(0.000362)
Constant	-1,403***	1.454***	-84.43***	-2.725***
	(48.64)	(0.0190)	(0.727)	(0.0302)
ICD9 Dummy Variables	Yes	Yes	Yes	Yes
Observations	556,201	556,201	556,201	556,201
R-squared	0.039	0.353	0.163	0.197

Interestingly, those who ADDED to their medical policy spent UAE dirham 163 more than those who remained in LOW plans and 123 more than those who stayed in HIGH plans. Notably, the prevalence of adverse selection as a result of adding to a current policy has never been explored in the Takaful insurance markets. Our results suggest that adverse selection is more prevalent in insurance markets where more PHs add to a medical policy than in markets with more/less switching, thus revealing higher switching costs. Therefore, additional research is needed to deepen our understanding of the effect of additional costs on the prevalence of adverse selection.

The results are unchanged when we consider the natural log of total expenditures as the dependent variable (Column 2). Switchers (UP and DOWN) spent 68 percent and 26 percent more than those who remained in LOW plans, respectively. Those who added to their current policy spent 50 percent more than those who stayed in LOW plans. Our results show that the extent of switching costs is larger than that found by Cardon (2018). For instance, he shows that those who switched up spent 23 percent more than those who stayed HIGH and that those who switched DOWN spent 4.7 percent more than those who stayed LOW.

When we consider the number of claims and the natural log of the claim number as dependent variables (Columns 3 and 4, respectively), the results are still consistent with our expectations. For instance, those who switched up filed six more claims than those who stayed LOW. However, those who switched down filed 16 percent fewer claims than those who stayed LOW. These results suggest that while they filed fewer claims, their average claim size was larger. This is very important, as a large number of claims filed does not necessarily indicate a higher risk. For instance, a healthy person may file 25 claims over a year with an average claim size of USD50, while an unhealthy person may file five claims with an average claim size of USD2,000.

Most of our control variables are significant at the 1 percent level across all specifications, and their predicted directions are consistent with the literature and our expectations. For example, as one would expect, having a chronic condition and being older are significantly associated with higher levels of medical expenditure. Moreover, consistent with Cardon (2018) and Eling, Jia, and Yao (2015), dependents (spouses and children) tend to utilize more insurance and claim higher amounts. However, the results for married people are inconsistent with Cardon (2018) and Eling, Jia, and Yao (2015), who find that married individuals spend less on medical insurance. Married individuals, however, tend to file more claims, suggesting that the size of their claims is lower. Males file fewer claims than females, and there is significant evidence of a difference in the amount of medical expenditure between genders.

Interestingly, PHs who had inpatient visits filed fewer claims than those who had outpatient visits. However, their medical expenditure was significantly larger (UAE dirham 399 more). This makes sense, as inpatient facilities provide more intensive medical treatments; thus, the size of the filed claims tends to be significantly larger. Furthermore, policy length is associated with more claims filed and larger medical expenditures. Finally, most ICD9 chapters (disease type) are associated with more claims filed and larger medical expenditures (the results are available upon request).

4.4.1 Additional Analysis and Robustness Tests

To show that our main results—the positive relationship between switching costs and the prevalence of adverse selection—are robust, we follow Finkelstein and McGarry (2006) and estimate the reduced-form models specified as Equations 4.2 and 4.3 in the preceding section.

Table 4.5 reports the results of estimating Equations 2 and 3. To conserve more space, we report only the results of our key variables. The results reported in Column 1 (Log Insurance Coverage Amount), Column 2 (Total Expenditure), and Column 3 (Total Number of Claims) confirm our findings reported in Table 4.4. Table 4.5 reveals that adverse selection is present in the market, represented by the positive relationship between insurance coverage in Column (1) and the dependent variables in Columns (2) and (3). Most importantly, the evidence of adverse selection is more evident in the positive correlation between the residuals ϵ and γ in Equations 2 and 3.

Regarding our key variable coefficients, the asymmetry of medical utilization found in Table 4.4 between those who add to their policy, nonswitchers, and switchers is robust. Furthermore, the differences between switchers and nonswitchers are significant, both economically and statistically.

Table 4.5. Robustness Test: The Effect of Switching Costs on Adverse Selection

This table is a robustness test using two reduced models (described in the methodology section) to confirm the effect of switching costs on the presence of adverse selection. The positive correlation between the residuals ϵ and γ in Columns 2 and 3 and Column 1 is a necessary condition for the presence of adverse selection. The key variables of interest are as follows: (1) switch up is a binary variable with a value of one if the policyholder (PH) switched up to a more generous plan; (2) switch down is a binary variable with a value of one if the PH switched down to a basic (economy) plan; (3) addition is a binary variable with a value of one if the PH added to a current policy; (4) stay high is a binary variable with a value of one if the PH remained in a generous medical plan; and (5) stay low is a binary variable with a value of one if the PH remained in a basic (economy) medical plan (staying low is the base). We include age, gender, marital status, relationship to the principal, policy length, chronic condition, visit type, and disease type (ICD9 categories) as controls in all regressions. Control variables are included and available upon request; refer to Table 4.1. Standard errors are reported in parentheses, and *, **, and *** indicate significance at the 10%, 5% and 0.1% levels, respectively.

	(1)	(2)	(3)
Variable	log(Coverage)	Total Exp.	Number of Claims
Stay High	1.271***	172.6***	7.938***
	(0.00197)	(4.022)	(0.0802)
Addition	0.766***	243.5***	5.499***
	(0.00210)	(4.295)	(0.0856)
Switch Down	0.673***	171.4***	3.696***
	(0.00247)	(5.041)	(0.100)
Switch Up	0.914***	358.7***	11.46***
	(0.00484)	(9.893)	(0.197)
Control Variables	Yes	Yes	Yes
Constant	12.28***	-120.6***	-13.62***
	(0.00579)	(11.82)	(0.236)
Observations	556,201	556,201	556,201
R-squared	0.571	0.036	0.146

4.5 Discussion and Conclusion

In a market with asymmetric information, switching between medical plans may cause a reduction in consumer welfare and exacerbate the presence of adverse selection. Handel (2013) provides empirical evidence that reducing inertia increases the overall switching propensities of opportunistic switchers at the expense of nonswitchers. In this paper, we empirically examine how switching between medical plans interacts with adverse selection in the context of healthcare insurance plans. In contrast to prior research, we further consider the effect of adding to a current medical policy (instead of switching UP or DOWN) on the extent of adverse selection. Our work controls for the presence of moral hazard by construction because we compare medical utilization between switchers and nonswitchers within the same medical plans (Cardon, 2018).

We provide significant evidence that adverse selection is preset in the market and that switching costs implies asymmetry of medical utilization among switchers, nonswitchers and those who add to their policy. Specifically, we show that the post-medical utilization of switchers, especially those who ADD to a current medical policy, was economically larger than that of nonswitchers who remained in HIGH or LOW medical plans. This suggests that adding to a current medical policy and switching UP to more generous medical plans are riskier than remaining in HIGH and LOW medical plans. It also suggests that the presence of adverse selection and/or asymmetric information is more pronounced among switchers than among nonswitchers and is even more pronounced among those who ADD/SWITCH UP than among those who switch DOWN and stay in LOW/HIGH medical plans. Cardon (2018) argues that "… such medical asymmetry is not simple adverse selection, but instead only appears because of the interaction of asymmetric information and significant switching costs. Therefore, the asymmetry offers a test for the presence of both adverse selection and switching costs."

References

Altman, D., Cutler, D. M., & Zeckhauser, R. J. (1998). Adverse selection and adverse retention. *The American Economic Review*, 88(2), 122-126.

Acerola, G. A., (1970). The Market for "Lemons": Quality Uncertainty and the Market Mechanism. *Quarterly Journal of Economics* 84(3), 488-500.

Cardon, J. H. (2018). Loss aversion and health insurance plan switching. *Journal of Economic Behavior & Organization*.

Cabral, M. (2016). Claim timing and ex post adverse selection. *The Review of Economic Studies*, *84*(1), 1-44.

Cardon, J. H., & Hendel, I. (2001). Asymmetric information in health insurance: evidence from the National Medical Expenditure Survey. *RAND Journal of Economics*, 408-427.

Cohen, A., & Siegelman, P. (2010). Testing for adverse selection in insurance markets. *Journal of Risk and Insurance*, 77(1), 39-84.

Cutler, D., Lincoln, B., & Zeckhauser, R. (2010). Selection stories: understanding movement across health plans. *Journal of Health Economics*, 29(6), 821-838.

Cutler, D. M., & Zeckhauser, R. J. (2000). The anatomy of health insurance. In *Handbook of Health Economics* (Vol. 1, pp. 563-643). Elsevier.

Deri, C. (2005). Social networks and health service utilization. *Journal of Health Economics*, 24(6), 1076-1107.

Einav, L., Finkelstein, A., & Levin, J. (2010). Beyond testing: Empirical models of insurance markets. Annu. Rev. Econ., 2(1), 311-336.

Eling, M., Jia, R., & Yao, Y. (2015). Between-Group Adverse Selection: Evidence from Group Critical Illness Insurance. *The Journal of Risk and Insurance*. 9999(9999), 1–39.

Handel, B. R. (2013). Adverse selection and inertia in health insurance markets: When nudging hurts. *American Economic Review*, 103(7), 2643-82.

Panhans, M. (2019). Adverse selection in ACA exchange markets: evidence from Colorado. *American Economic Journal: Applied Economics*, 11(2): 1–36

Polyakova, M. (2016). Regulation of insurance with adverse selection and switching costs: Evidence from Medicare Part D. *American Economic Journal: Applied Economics*, 8(3), 165-95.

Robinson, J. C., Gardner, L. B., & Luft, H. S. (1993). Health plan switching in anticipation of increased medical care utilization. *Medical Care*, 43-51.

Rothshild, M., and Stiglitz, J. (1976). Equilibrium in Competitive Insurance Market: An Essay on the Economics of Imperfect Information. *Quarterly Journal of Economics*, 90(4), 629-647.

Rowell, D., Nghiem, S., & Connelly, L. B. (2017). Two tests for ex ante moral hazard in a market for automobile insurance. *Journal of Risk and Insurance*, 84(4), 1103-1126.

Simon, K. I. (2005). Adverse selection in health insurance markets? Evidence from state small-group health insurance reforms. *Journal of Public Economics*, 89(9-10), 1865-1877.

Tchernis, R., Normand, S. L. T., Pakes, J., Gaccione, P., & Newhouse, J. P. (2006). Selection and plan switching behavior. INQUIRY: *The Journal of Health Care Organization, Provision, and Financing*, 43(1), 10-22.

5. Chapter Five: Conclusion

5.1 Thesis Review

In this thesis, I contribute to the existing literature on insurance and health economics through three related studies.

The **first study** focuses on how culture and religion influence asymmetric information in the Takaful healthcare insurance market. We provide significant evidence of asymmetric information in the Takaful health insurance market. We also find evidence that religious individuals are less likely to file claims and file for claims at smaller amounts than nonreligious individuals. Additionally, we find some evidence consistent with our expectation that cultural background is associated with claims behavior and choice in insurance coverage. Finally, we develop a theoretical model to capture the role of religiosity on Takaful insurance demand and its effect on the negative drivers of asymmetric information (e.g., fraud, misrepresentation, and withheld information). In this simple model, we show that when holding other factors constant, religiosity (as an intrinsic value driver) is a key factor in Takaful insurance demand for religious Muslims and can be a significant driver in reducing negative behaviors that influence the presence of asymmetric information. We show that the utility functions of irreligious and non-Muslim policyholders emphasize the price differences in comparing Takaful and commercial insurance products and thus that they would demand Takaful insurance products only if the Takaful products' coverage terms are superior to those of the commercial insurance counterparts, since the price of Takaful insurance is higher.

The **second study** examines whether and to what extent fraud is committed largely by HCPs rather than by healthcare insurance PHs. We then apply principal component analysis of RIDIT scores (PRIDIT) to further deepen our understanding of healthcare insurance fraud and identify the variables and claims with the strongest indicators of suspicion. Our results show that fraud/abuse activities in the healthcare insurance market are committed mostly by HCPs rather than by PHs. Furthermore, fraudulent claims are more likely to be rejected when the HCP rather than the PH controls the claim.

The **third study** examines how switching between medical plans interacts with adverse selection in the context of healthcare insurance plans. In contrast to prior research, in our study, we further consider the effect of adding to a current medical policy (instead of switching UP or DOWN) on the extent of adverse selection. Our work controls for the presence of moral hazard by construction because we compare medical utilization between switchers and nonswitchers within the same medical plans (Cardon, 2018). We provide important evidence that adverse selection is preset in the market and that switching costs implies asymmetry of medical utilization among switchers, nonswitchers and those who add to their policy. Specifically, we show that the post-medical utilization of switchers, especially those who ADD to a current medical policy, was economically larger than that of nonswitchers who remained in HIGH or LOW medical plans. Our results suggest that adding to a current medical policy and switching UP to more generous medical plans are riskier than remaining in HIGH and LOW medical plans. It also suggests that the presence of adverse selection is more pronounced among switchers than among nonswitchers and is even more pronounced among those who ADD/SWITCH UP than among those who switch DOWN and stay in LOW/HIGH medical plans.

5.2 Future Research

Economists have examined the effects of religion and culture on economic behavior (e.g., Iyer, 2016; Iannaccone, 1998; Hilary and Hui, 2009; Guiso, Sapienza, and Zingales, 2003). However, the empirical work within the insurance literature, which is the focus of our study, is not well documented. Insurance economists have examined how religion, as a proxy for national culture, influences insurance demand across various countries, including some Muslim countries (e.g., Outreville 2018, Trinh, Nguyen, and Sgro 2016, and Park and Lemaire 2012). Although we have highlighted the influence of religiosity on Takaful insurance demand and asymmetric information, our research raises additional questions regarding the extent to which this demand and asymmetric information change if we relax some of the assumptions (e.g., risk factors, initial wealth, and Takaful-commercial prices differential) from our simple model. For instance, what if Takaful insurance prices experience a significant reduction or increase? Furthermore, is there a relationship between Takaful insurance demand and the level of asymmetric information present? The determination of the rationale behind Takaful insurance demand, given a policyholder's religion and other economic preferences, and its potential effect on negative economic behaviors is important for understanding individual insurance decisions and is thus worth further investigation. We encourage insurance economists to further explore the roles of religion and culture on Takaful insurance demand and their effects on the prevalence of asymmetric information and fraud across all insurance markets (e.g., Takaful vs. commercial). This section serves as a cornerstone for future extensions and should help researchers to develop advanced theoretical models to better understand the role of religion on asymmetric information and Takaful insurance demand. Our efforts also resonate with the call from Schlesinger (2013) who is stating that "…I look forward to seeing the directions in which the theory of insurance demand is expanded in the years to come, and am encouraged to know that some of you who are reading this chapter will be playing a role in this development."

Additionally, economists have already examined the effect of medical plan switching on adverse selection and have argued that switching between medical plans predicts future utilization (e.g., Cardon 2018, Handel 2013, Cutler et al. 2010, Tchernis et al. 2006, Cardon and Hendel 2001, Robinson et al. 1993). Our results in **Chapter 4** before and after switching up and down are consistent with the current literature, especially Cardon (2018), who examines the effect of loss aversion on medical plan switching behavior. However, the prevalence of adverse selection pre- and post-addition (empirically and theoretically) is not well documented by insurance economists. Thus, we encourage insurance economists to further explore the potential effect of switching behaviors on the extent of adverse selection, especially among those who add to their policy instead of switching up or down.

Finally, empirical research on insurance fraud is still limited due to the lack of data, especially in the healthcare insurance markets (e.g., Ai et al 2018, Tennyson and Salas-Forn 2002, and Derrig 2002). However, insurance economists have been focusing on public and social insurance markets (e.g., Fang and Gong 2017, and Pande and Maas 2013). Most of the current analysis focuses on understanding the relationships between different predictor variables and PRIDIT (fraud) scores (Brockett et al. 2002, Ai et al. 2003, and Golden et al. 2019). Although we identify in **Chapter 3** the most suspicious claims in our dataset, we encourage insurance economists to relate them in future studies more directly to the cost of processing claims, settlement time, and claim disposition, as further analysis is required to identify the reasons for rejection as well as to conduct a more granular analysis of the specific types of medical services that result in rejected claims. In addition, to refine our results further, we encourage insurance economists to divide the analysis by ICD9 code, HCP type (e.g., doctor, emergency center, and lab), and Hofstede cultural variables (e.g., individualism, power distance, masculinity, and uncertainty) and relate them to healthcare insurance fraud. We plan to conduct this additional analysis, which is beyond the scope of this thesis, in future studies.