Sustainability and Adverse Selection in Emerging Health Insurance Markets: Evidence from Microinsurance in Pakistan

By

Yi (Kitty) Yao

A dissertation submitted in partial fulfillment

of the requirements for the degree

of Doctor of Philosophy

(Business)

at the

UNIVERSITY OF WISCONSIN-MADISON

2012

Date of final oral examination: 07/02/12

The dissertation is approved by the following members of the Final Oral Committee:

Joan T. Schmit, Professor, Actuarial Science, Risk Management and Insurance

Justin R. Sydnor, Assistant Professor, Actuarial Science, Risk Management and Insurance

Mark J. Browne, Professor, Actuarial Science, Risk Management and Insurance

Edward W. Frees, Professor, Actuarial Science, Risk Management and Insurance

John Mullahy, Professor, Population Health

Dedication

This dissertation is gratefully dedicated to my beloved parents, Jianling Xu and Guang Yao, and to my dear husband, Rui Wang.

Acknowledgements

Finishing a dissertation, as the last and most challenging step in pursuing a Ph.D., could be a lonely, painful, and frustrated undertaking that lasts for years and seems endless. It is said, “Rome wasn’t built in a day,” but then the glorious day comes when someone finally says “Congratulation! You’re done.” When that day came, I looked back on these past five years and realized it has been an amazing journey that has taken me to the place of my dreams. I made it to this point because I was surrounded by so many extraordinary people who made this journey so inspiring, exciting, and joyful.

First and foremost, I want to thank my parents, Jianling Xu and Guang Yao. Gaining a Ph.D. is a family investment, both monetarily and emotionally. It must have been difficult for them to let their only child go halfway around the world from them. It made the clock tick slower. Although they have occasionally complained about this seemingly endless five years, they have been truly the most supportive and loving parents imaginable. I know that completing this degree will bring happiness to everyone in my family. If my grandpa were alive, he would still be sitting in his small village community center, playing chess with his friends, and bragging about his granddaughter studying in America for something that he never understood but nevertheless made him very proud. The expectations of my family have been the force driving me forward.

Next, I thank my advisor, Joan Schmit, and my co-advisor, Justin Sydnor. Joan is amazing. I don’t have enough space to describe all the memorable things about her that flash through my mind, because I truly could write a dissertation about her, one that probably would be much easier than the one that I have worked on. Serving as associate dean and interim dean with a desk piled high with files that "required a tornado to clear" (in her words), she has had every reason to be a distant advisor; however, she has been just the opposite. She is an inspiring teacher, a devoted mentor, a dear friend, an amiable person, an effective communicator, an elegant woman, and a role model for me in years to come. Choosing Justin as my co-advisor was one of the best decisions that Joan suggested for me. He is smart and generous, always willing to share his experience and offer help. Those Wednesday afternoon talks in his office were the start of a transition for me, shaping me toward thinking as a young scholar instead of as a student.

I followed a zigzag route in pursuing my data sources, because a standardized commercial dataset of microinsurance is still nonexistent. Numerous people (many of whom I have never met) guided me through this process. These include Jinwen Cao of the China Life Asset Management Corporation, Zhigang Zhai of China Life, Benxin Yang of Xinhua Life, Michael McCord, president of the Microinsurance Center, Denis Garand, president of DGA, and Mary Yang of the Microinsurance Innovation Facility of the International Labour Organization. These amazing contacts who embraced the interest of a total stranger guided me to the Aga Khan Agency of Microfinance, where I finally obtained the dataset used in this analysis. Special thanks go to Peter Wrede, AKAM's former microinsurance specialist and now an officer in Microinsurance Innovation Facility of the International Labour Organization. I'm so impressed by Peter's passion for microinsurance. He is generous, patient, responsive, professional, and warmhearted. I have never been to Pakistan to investigate the microinsurance project myself, and Peter was of tremendous help, via numerous communications through phone, Skype, and e-mail, to my understanding of the background and operation of microinsurance there.

I thank my husband, Rui Wang, from the bottom of my heart. He is the best “side effect” of my pursuit of a Ph.D.. We were classmates and are now committed “life mates.” Despite all the ups and downs of our life, he is firm, strong, always positive and supportive. I could not imagine going through all these years without him by my side. He was usually the first discussant of my research ideas, the first reader of my drafts, and the first audience for my presentations. He has asserted that he knows more about my dissertation than anyone on my committee, and that’s probably true.

Last but definitely not least, I thank all my other committee members — Mark Browne, Jed Frees, and John Mullahy — and faculty members Margie Rosenberg, James Guszcza, and Martin Halek. I am also deeply appreciative of all the doctoral students in my department who have been on this journey with me: Lan Ju, Peng Shi, Shinichi Kamiya, Winnie Sun, Cuncun Luan, Chunyan Zhang, Xiaoli Jin, Joyce Lin, Marc Ragin and Marc-Andre Desrosiers. All of them helped me in different ways, and all together we have made the department feel like home. I am also grateful to Shiwei Qumu (M.D.) for her professional consultations regarding healthcare claim data categorization. And I owe a great deal of gratitude to my friends and former roommates in Madison for filling the long winters with warmth and laughter.

All of you are the wind beneath my wings.

Abstract

Despite widespread interest in expanding insurance at the bottom of the economic pyramid, the viability of emerging microinsurance programs has been questioned because of high loss ratios and dubious sustainability. The purpose of this research is to derive implications for the provision of viable products for emerging microinsurers by investigating the degree of sustainability of a micro health insurance program in its early years of development and determining whether adverse selection exists within such programs.

To investigate sustainability, I used data from a micro health insurance program in Pakistan to analyze how claim rates evolve as households renew their policies. I found that households with larger claims during the initial policy year are more likely to renew their policies. Although that pattern appears consistent with decreasing sustainability, I found in delving deeper that when compared with households buying insurance for the first time, renewed households have significantly lower frequency of claims and also lower total claim amounts.

In addition, to determine if adverse selection existed, I further categorized the claim data from the same microinsurance program. Compared to households that filed only claims for acute diseases, those that filed claims for chronic diseases are more likely to renew their policy. A prevalence of adverse selection is observed within maternity-related claims.

The combination of these results shows that despite the existence of adverse selection and the fact that households with claims are more likely to renew their policies, renewed households have a much improved claim frequency and total claim amount after their first year of coverage.

This could be explained by the prevalence of pre-existing conditions among newly enrolled households that are encountering affordable healthcare for the first time. In addition, some selection issues, such as pregnant women selecting to enroll in the program also contribute to the high first-year claim frequency among newly enrolled households. These factors taken together explain why the loss ratio for the renewed households actually decreases dramatically after the first year of coverage.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Background

Although low-income individuals in developing countries rank health insurance as one of their most needed forms of insurance, little such coverage is available to them because of the complexity of delivering health insurance in less developed areas. Both acute and chronic diseases are serious problems for the poor in these countries, where poor sanitation, poor nutrition, and inadequate preventive care are widespread. As a result, the life expectancy for those living in low- and middle-income countries is 20% lower than for persons in high-income countries. In an extreme comparison, the life expectancy for people living in sub-Saharan Africa is only 46 years compared to 75.5 years in high-income countries.[1] For the poor, the consequences of illness include not only reduced current living standards but also the potential to lead them to reduce investment in their future human capital. Against that backdrop, it is perhaps unsurprising that people with low incomes also have an interest in health insurance programs. Despite their interest, however, only 20% of low-income persons are believed to have access to adequate health insurance.[2]

In recent years interest in the creation of new insurance products in developing countries has grown dramatically. In addition to the involvement of nonprofit organizations and governments in the financing and operating of these programs, more and more commercial

---

[1] Table 2.3. Selected mortality characteristics by sex and World Bank region in 2001, Global burden of disease and risk factors, Oxford University Press and The World Bank, 2006.
[2] Bockstal (2008)

insurance companies have started to step in. At least 33 of the 50 largest commercial insurance companies in the world offered microinsurance products in 2012, up from only seven in 2005, according to the Microinsurance Innovation Facility of the International Labour Organization and the Munich Re Foundation.[3] Moreover, the number of policyholders covered by these programs increased substantially from 2007 to 2012. Estimates are that the number increased from 78 million in 2007 to 135 million in 2009, reaching 500 million in 2012, a sixfold increase in five years.

From the start, interest in these new "microinsurance" products has been closely related to the growth of interest in "microfinance," serving as a "byproduct" of microfinance product packages in the form of voluntary or compulsory credit life insurance. Over time, microinsurance turned to be more versatile, becoming available across almost all lines of business and taking on various organizational forms and delivery models. To be more specific, micro health insurance follows four types of delivery models: partner-agent, provider-driven, full-service, and community-based.[4] The primary differences among these models lie in who is in charge of program design and development as well as in the relationship of the healthcare provider and/or policyholders to the management of the microinsurance program.

No matter which specific business model a microinsurer chooses, the choice reflects efforts to find a way to provide insurance to the low-income segment of the developing world. To achieve this goal, however, the microinsurer must find a sustainable business model. Emerging economies face particular challenges in developing insurance programs. The reasons for these challenges include relatively high expense ratios, given the lower face values relevant to the

[3] News release, "Microinsurance coverage expanding at breathtaking pace according to ILO and the Munich Re Foundation," (April 10' 2012)
[4] Churchill (2006)

insuring population; the higher probabilities of loss because of poorer health; limited infrastructure; more dangerous living and working conditions; and lack of understanding of the insurance mechanism. Therefore, the sustainability of any microinsurance mechanism in emerging economies becomes questionable.

For a variety of reasons, sustainability has been an issue for many microinsurance schemes already attempted. From the funding side, many programs start with significant subsidies from governmental or nonprofit organizations; however, in most cases, this arrangement lasts only for a specified number of years after which the insurer must be able to stand on its own. On the business side, many programs experience high loss ratios and high lapse rates in their start-up years, leading generally to financial losses. Other researchers have suggested that poor underwriting experiences are owed to worsening claim experiences created when only poor risks decide to stay with an insurer. My intention is to contribute to an understanding of the development of such important insurance mechanisms in emerging economies.

1.2 Objective

The purpose of this research is to derive implications for the provision of viable products for emerging microinsurers by investigating the degree of sustainability of, and the impact of any adverse selection on, a micro health insurance program in its early years of development.

Although the long-term financial success of micro health insurance programs requires sound data analysis, limited empirical knowledge currently exists about this market. The scant literature available conveys conflicting views about the potential for sustainable micro health insurance. Pauly et al.[5] used data from the World Health Survey of 14 developing countries to compare the risk premium, with likely values for administrative expenses that local people would be willing to pay. Without taking information asymmetry into account, they concluded that a voluntary health insurance market might be feasible. In contrast, Biener and Eling[6] examined the problems and solutions for microinsurance markets. Based on a survey of studies, they concluded that problems of information asymmetry were "epidemic" in micro health insurance and cited as evidence examples in the literature that ranged from adverse selection to moral hazard to fraud, all of which generated concerns over sustainability.

In this paper, I address this debate in two ways. First, I focus on the development of a microinsurer's risk portfolio in the insurer's early years of operation. Specifically, I examined households' renewal decisions and the results reveal the claim trends in renewal business. A better understanding of these trends and of the forces driving them will help improve assessment of the sustainability of micro health insurance programs. Second, I examine claims in different risk categories and how these risk categories affect households' renewal decisions. The goal is to be able to speak to adverse selection and link it to sustainability.

1.3 Outline

The remainder of this paper proceeds as follows. Chapter 2 reviews the literature on sustainability and information asymmetry in general and explores the small body of literature on micro health insurance. Chapter 3 introduces the Aga Khan Agency of Microfinance (AKAM) micro health insurance program and the data used in the empirical analysis. Then Chapter 4

[5] Pauly et al. (2008)
[6] Biener and Eling (2012)

focuses on models exploring the development of claim trends in renewal business, and Chapter 5 discusses the models used to present evidence concerning adverse selection.

Chapter 2

Studies on Sustainability and Adverse Selection

2.1 Sustainability of the Micro Health Insurance Market

According to Churchill and Garand,[7] the main strategy that the industry used to achieve sustainability of insurance products consists of three elements, namely limiting benefits, focusing on efficiency, and diversifying income sources.[8] Most academic work, however, has focused mainly on ways to improve efficiency in a given framework of available resources and benefits promised.[9]

In general, measuring and evaluating the performance of microinsurance programs takes one of two approaches. One approach uses performance indicators as transparent and comparable measurements that microinsurance practitioners can use as benchmarks.[10] The Consultative Group to Assist the Poor (CGAP) microinsurance working group (now known as the Microinsurance Network) has developed eight key principles and ten key financial ratios to use as guidelines to measure microinsurers' performance. These indicators cover ratios for net income, incurred expenses, incurred claims, renewal rates, solvency, claim rejection, growth, coverage, liquidity, and a measure of promptness of claim settlement. Because serving low-income people is a fundamental goal of microinsurance, financial performance measures alone are insufficient. Therefore, the CGAP working group on microinsurance also considers four

[7] Churchill and Garand (2006) Chapter 6.1 Strategies for Sustainability, Protecting the Poor: A Microinsurance Compendium, edited by Craig Churchill, Munich Re Foundation and International Labour Office, p564
[8] The Consultative Group to Assist the Poor (CGAP) working group on microinsurance (2007)
[9] The Consultative Group to Assist the Poor (CGAP) working group on microinsurance (2006), Biener and Eling (2011), Zheng and Zhang (2010)
[10] The Consultative Group to Assist the Poor (CGAP) working group on microinsurance (2006)

potential social measures, including the social investment ratio (defined as total expenditure on information, education, and communication divided by the total expenditure of the program), the percentage of insureds below the poverty line, the value of incurred claims in comparison to client annual income, and the cost of benefits provided in comparison with the annual premium.[11]

The goal of this first approach is to create all-around practical performance measurements using standard financial ratios as well as some social measures, and the Microinsurance Network intends to build a database of these measurements for microinsurers to use for program design and comparison purposes. The second approach to measuring and comparing program efficiency relies on a more sophisticated method using Data Envelopment Analysis (DEA). DEA is a well-accepted and rapidly innovating nonparametric method to estimate an efficiency frontier and sets a benchmark for industry performance. Biener and Eling[12] used the DEA method to analyze samples from 2004 through 2008 that were drawn from 20 micro life and health insurance programs. Unlike the financial ratio method, the DEA approach is adapted from multi-input and multi-output production functions and summarizes the different characteristics of a firm in a single efficiency measurement. Biener and Eling found significant potential for improvement in the efficiency of microinsurance programs, especially those of small and nonprofit microinsurers. They found the sale of group policies to be more efficient than selling individual policies and recommended adoption of this more efficient sales approach to reduce transaction costs and also overcome information asymmetry problems. Similarly, Zheng and Zhang[13] also adopted the DEA method to evaluate the efficiency of China's New Rural Cooperative Medical System

[11] Biener and Eling (2011)
[12] Biener and Eling (2011)
[13] Zheng and Zhang (2010)

(NRCMS) in developing comparative data to measure the patterns of efficiency across eastern, central, and western districts of China.

When either of these performance measures indicates a microinsurance program is in trouble, one proposed solution has been to link it with a Micro Finance Institute (MFI) as a way to boost sustainability. Such a linkage results in a shared distribution channel that reaches potential customers at a lower cost. This proposal has merit, on the one hand, because microinsurance programs often originate from MFIs in the form of voluntary or compulsory credit life insurance for MFI customers. The proposed combined business model manages risk for both the MFI and its policyholders; it also opens the way for microinsurance programs to grow rapidly and gain access to a vast number of potential customers while minimizing administrative and distribution costs.

On the other hand, an argument can be made that combining with the MFI also benefits a microinsurance program via low-cost risk underwriting. For example, to get a loan approved, applicants must demonstrate an ability to do business and come to the loan office to apply. Therefore, this group of potential customers could be a relatively good risk as purchasers of micro health insurance. The microinsurance program could benefit from this low-cost underwriting mechanism.

This linked relationship, however, also has its drawbacks. The most important one is that the microinsurer and the MFI may have different goals and interests. Therefore, a combined relationship with an MFI is a suitable source for certain kinds of microinsurance products, but not for all of them.

To illustrate, Kwon[14] found that certain organizational, market, and sociocultural factors affected the willingness of MFIs to expand into microinsurance service. He used data from over 600 MFIs that were operating from 1998 to 2007 in 83 countries and found that MFIs in business the longest and with larger customer volume, larger families, and higher ratios of both women customers and life insurance penetration were more likely to offer microinsurance products.

The research presented in this paper uses neither the CGAP performance benchmark nor DEA method to assess sustainability. Those two general approaches work best when comparing efficiency across multiple microinsurance programs, but in this paper I focused on evaluating the sustainability of one particular microinsurance program. My focus is on AKAM program, which operated a micro health insurance branch in Pakistan from 2007 through 2011.[15] Instead of comparing the performance of this program with some industry benchmarks or similar programs, I examined the sustainability of this micro health insurance by comparing its own history and book of business through an analysis of households' renewal decisions. Specifically, I analyzed the trend of sustainability by measuring the development of claim experiences in renewal policies. I used a two-part model to compare the microinsurer's new book of business to its renewed book of business and to test whether the risk pool had deteriorated over time. The status of its risk pool over time sheds light on the sustainability of a microinsurer's business.

Asymmetric learning and aging phenomenon often play key roles in the sustainability of an insurance program. Both conditions affect an insurer's profit trend over time as policyholders choose to renew or drop their policies. Asymmetric learning in an insurance market refers to the results when an incumbent insurance company knows more than its competitors about the risks

[14] Kwon (2010)
[15] In April 2011, the program was moved from AKAM to the New Jubilee Life Insurance Company.

associated with its customer base and therefore is able to assess risk more accurately and thus improve its profits over time.

The aging phenomenon is a piece of asymmetric learning and refers to the trend that insurers, especially in property-liability lines of businesses, typically earn greater profits on policies with renewed policyholders than on newer business. Kunreuther and Pauly as well as D'Arcy and Doherty[16] postulate that this is caused by insurers' accumulation of private information. Over time, the insurer learns the true type of risk that a policyholder represents thus is able to adjust the price for the next policy year accordingly. As a result, high-risk policyholders will get a premium increase, driving them to a new insurer. The new insurer lacks private information about the policyholders' true type of risk and consequently levies a lower premium. As a result, the original insurer assembles a better risk portfolio of renewing business and reaches a lower loss ratio compared with its new business.

Kofman and Nini[17] tested three predictions of asymmetric learning. First, insurance claims and consumer switching should be contemporaneously positively correlated. Second, the average risk associated with a cohort of insured should decrease over time; and third, the average profit associated with a cohort of insureds should increase over time. Data from a large Australian insurance company, however, failed to support the second and third predictions. Moreover, average risk and profitability were shown to be constant over time, suggesting that insurers did not possess an information advantage over their competitors.

Although on the surface the relationship tested in this study looks very similar to the aging phenomenon and asymmetric learning, the underlying mechanism differs. The aging

[16] Kunreuther and Pauly (1985); D'Arcy and Doherty (1990)
[17] Kofman and Nini (2004)

phenomenon is a trend in property-liability lines in a developed insurance market that shows an insurer retained a lower loss ratio in renewed business than in new business. The underlying requirements for the aging phenomenon to occur are twofold. First, many insurance providers must be available for policyholders to switch to; second, insurers must be enforcing an experience rating based on policyholders' previous claim histories. These two requirement, however, are absent in the AKAM micro health insurance program because it is the only health insurance provider in the area and it did not adjust its premiums based on claim history.

2.2 Adverse Selection in the Micro Health Insurance Market

2.2.1 Information Asymmetry: Review of Theory and Existing Approaches

Adverse selection and moral hazard are two consequences of having asymmetric information between a principle and an agent. In the context of the insurance market, it usually refers to the fact that an insured has private information that an insurance company cannot observe and therefore, cannot take into account in pricing. This could jeopardize the insurance market and lead it to a "death spiral" in an extreme case, as described in Akerlof.[18] Asymmetric information differs from asymmetric learning. The former emphasizes asymmetric informational differences between insurer and policyholder, and the latter focuses on differences in information acquisition between a contracting insurer and other insurers in regard to information about policyholders.

[18] Akerlof (1970)

The theory of asymmetric information was first established in the 1970s through the seminal work of Akerlof, Pauly, and Rothchild and Stiglitz,[19]and further developed by later generations of scholars.[20] In the classic Rothchild-Stiglitz model, insureds were assumed to be of different risk types, and the model described the market equilibrium under information asymmetry. Under the assumption that the insurer cannot distinguish high risk from low risk in the market, a pooling equilibrium with all the insureds paying the same level of premium will not be sustainable because the low-risk insureds will gradually drop out, a case similar to the lemon market that Akerlof described. Even in the separating equilibrium in which high risks buy full insurance coverage and low risks buy partial coverage, the low risks will suffer a welfare loss. Miyazaki and Wilson further developed the model by including additional assumptions concerning nonnegative profit and expectations of interaction among firms.

These theories helped to specify risk classification as a solution for adverse selection because if an insurance company could differentiate between low-risk and high-risk purchasers, it could sell different contracts to both groups and improve the welfare of society. In addition, these theories established the rationale for testing for adverse selection in a given market by using the "positive correlation test" between risk type and insurance coverage purchased. For example, Puelz and Snow[21] used data from the auto insurance market of the U.S. and found those with higher accident risks chose lower deductibles (more insurance coverage), a finding that agreed with the predicted outcome of the "positive correlation test."

---

[19] Akerlof (1970); Pauly (1974); Rothchild and Stiglitz (1976)
[20] Miyazaki (1977); Wilson (1977); Finkelstein and McGarry (2006)
[21] Puelz and Snow (1994)

Scholars have tested for adverse selection in various markets by using different proxies for risk type and insurance coverage, and the empirical results are mixed.[22] The most common proxies for insurance coverage are policies with different levels of deductibles and copayments, decisions to opt in and out of insurance, and the option to purchase supplemental insurance. Proxies for risk type range from subjective measurement (self-evaluated health condition) to objective ones (indicators such as age and medical history) and predicted risk type.[23]

Browne[24] used data from the individual health insurance market in the U.S. to test for adverse selection, finding that low-risk policyholders purchase less insurance in the individual health insurance market than they would have in a group market; this suggests the existence of adverse selection in the individual health insurance market. Another work by Browne[25] examined the long-term care market in the U.S. by using data from one large insurer. He found insureds with high predicted loss amounts were more likely to retain coverage in the following year despite an annual premium increase, again suggesting the existence of adverse selection.

Using data from the auto insurance market, Cohen[26] tested the predictions of adverse selection models among newly enrolled policyholders who were assumed to have an informational advantage over the insurance companies. She found that new customers who chose lower deductibles, i.e., higher insurance coverage, were associated with more accidents and higher total claim amounts. This correlation is especially significant for drivers with three or more years of driving experience, but not significant for new drivers with limited experience.

[22] See Cohen and Siegelman (2010) for a review.
[23] Browne (1992, 2006); Gao et al. (2009)
[24] Browne (1992)
[25] Browne (2006)
[26] Cohen (2005)

This suggested that the policyholders were learning about their true type of risk and then adjusting their selection of insurance coverage options accordingly.

In contrast, Dionne et al.[27] also used data from the auto insurance market to examine the "positive correlation test" and challenged Puelz and Snow's conclusion.[28] They used data from a large Quebec insurer to reproduce a similar empirical regression and argued that the Puelz and Snow results were not robust. Indeed, their policyholders with more claims were more likely to choose a higher deductible, which agrees with Puelz and Snow. But when controls for the expected number of claim were introduced into the regression, the correlation between the number of claims and the choice of higher deductibles was no longer significant. This result shows that an insurer is able to control for adverse selection by including estimates of the risk type of policyholders, and after that no significant residual adverse selection remains in the market.

Others[29] have found a negative relationship between risk type and insurance coverage, which indicates that low-risk individuals purchased more insurance coverage. This negative relationship observed in empirical tests was referred to as advantageous selection in contrast to the traditional term of adverse selection.

In contrast to the previous studies that focused solely on risk type, more recent researchers have extended the theory of adverse selection by showing that private information has multiple dimensions besides risk type. This broadened perspective explains the empirical finding of advantageous selection.

[27] Dionne et al. (2001)
[28] Puelz and Snow (1994)
[29] Bolhaar et al. (2008); Fang et al. (2008); Gao et al. (2009); Einav and Finkelstein (2011)

Chiappori and Salanie[30] examined auto insurance data from France and found no evidence of a significant correlation between the auto insurance coverage purchased and risk type (measured by accident history). One of their explanations is the "cherry picking theory" of risk preference, i.e., risk averse individuals will purchase more insurance and be involved in fewer accidents.

Finkelstein and McGarry [31] presented a model with multiple dimensions of private information, namely risk type and risk preference, which could act in opposite directions to offset the occurrence of risk. Using data from long-term care insurance in the U.S., they found the existence of private information, but the correlation between insurance coverage and risk occurrence was not significant after controlling for risk classification, which meshes with the theory of multiple dimensions of private information.

In addition to risk preference, various other sources of advantageous selection have been proposed, including heterogeneity in income, education, health preferences, financial planning horizons, and cognitive ability. A study conducted by Gao et.al[32] showed both adverse selection and advantageous selection in China's health insurance market, and they argued that both phenomena were consistent with the existence of information asymmetry. They showed that insureds who eventually filed more claims tended to have purchased lower limits of basic insurance, but were more likely to have purchased additional insurance. With a theoretical model analysis, they found that lower risk individuals may purchase more basic coverage because of the heterogeneity in wealth levels and loss amounts, thus exhibiting advantageous selection in the market.

---

[30] Chiappori and Salanie (2000)
[31] Finkelstein and McGarry (2006)
[32] Gao et.al. (2009)

Different sources of advantageous selection were further scrutinized in Fang et.al.[33] Using data from the Medigap insurance market in the U.S., they found that the sources of advantageous selection include income, education, longevity expectations, financial planning horizons, and cognitive ability. The latter, they found, is particularly important. However, in their study risk preferences do not appear as a source of advantageous selection.

2.2.2 Information Asymmetry in the Micro Health Insurance Market

Although the basic predictions of insurance theory should generally hold true for microinsurance, the nature of the product poses a number of challenges to offset the impact of adverse selection in a developing market. As Brau et al.[34] pointed out, the impact of adverse selection for a microinsurance carrier is critical because of the cost of gathering information regarding risk. Microinsurance as an emerging product for low-income people in developing countries could suffer from a perilous degree of adverse selection, especially because of a limited capability to assess and classify risks. Because of the need to reduce administrative costs while keeping the product simple, microinsurance products are often designed to be universal without deductibles or coinsurance options.

This design for universality is mainly for two reasons. First, it fits the financial literacy of low-income individuals. Usually a microinsurance policy is the first insurance policy a low-income individual purchases; deductibles and coinsurance require more sophisticated knowledge about insurance than these individuals usually possess. Without proper education of purchasers, and easily understood handling of their claims, misunderstandings could occur that would ruin

[33] Fang et.al (2008)
[34] Brau et al. (2011)

the reputation of the insurer and of insurance overall before the market could develop fully. Second, insurers are reluctant to implement more sophisticated policy features because they want to hold down administrative costs. Effectively administering policies with deductibles and coinsurance requires electronic recordkeeping, and it may be necessary to implement a system to track and verify claims onsite at every clinic. Most microinsurers have not yet invested in these technologies and infrastructure.

Because microinsurance markets remain relatively immature and data are often inaccessible and imperfect, only a few empirical studies have focused on testing for adverse selection in this market. Those few researchers who have studied this market with various methodologies and data from different markets in different time periods have found mixed evidence with regard to the existence of adverse selection.

A few studies support an argument for the existence of adverse selection in micro health insurance programs. For example, a series of studies on adverse selection have been conducted using data from a rural mutual healthcare insurance project in China. Wang et al.[35] followed a voluntary mutual healthcare insurance program there from 2002 to 2006 and performed a panel data analysis. Using individual level data, they found strong evidence of adverse selection despite a high enrollment rate of 71% and a requirement that an entire household enroll as a unit. In particular, they found that the pre-enrolled medical expenditures for enrolled individuals were 9.6% higher than the average expenditure for all residents. In addition, the inability to strictly enforce the requirement that an entire household has to enroll altogether resulted in partial enrollment of about a third of the enrolled households. The enrolled members of a partially enrolled family spent 1.7 times more than the non-enrolled members of a partially enrolled

[35] Wang et al. (2006)

family. They concluded that without taking adverse selection into full consideration, the voluntary program could not be sustainable financially.

A follow-up study of the same program by Zhang and Wang[36] also observed that people with histories of chronic conditions and those in fair or poor health overall were more likely to enroll in the program, thus demonstrating the existence of adverse selection. The authors also used a four-year longitudinal dataset to examine the effects of adverse selection over time by including interaction terms between health status and different wave variables. They found the extent of adverse selection seemed to be stable over the study period.

Using historical survey data of U.S. short-term disability microinsurance in the early twentieth century, Murray[37] found *prima facie* evidence of asymmetric information in which evidence of the presence of adverse selection outweighed the presence of moral hazard. In addition, he showed that the countermeasures taken by the microinsurers, including the enforcement of trial periods and waiting periods, effectively reduced claims. Similarly, Ito and Kono[38] found evidence of adverse selection in a micro health insurance program in India because households with a higher ratio of sick members were more likely to purchase insurance. Moreover, Lammers and Warmerdam[39] also examined adverse selection in voluntary micro health insurance. Their work used data from Nigeria and showed, after controls for risk preference and demographics, including wealth, that recent illness and self-prediction of health risk increased the probability of enrollment. Moreover, Clement[40] also found evidence for adverse selection in the Ghana national health insurance program, showing that high-risk

[36] Zhang and Wang (2008)
[37] Murray (2011)
[38] Ito and Kono (2010)
[39] Lammer and Warmerdam (2010)
[40] Clement (2009)

individuals were more likely to select into the program. In particular, Clement found that women in their child-bearing years were more likely to enroll, and most of them enrolled after they found they were pregnant. Also, adults older than 50 were more inclined to enroll.

Using survey data from seven micro health insurance programs in India, Dror *et al.*,[41] studied the determinants of willingness to pay for micro health insurance. They found evidence to support the existence of adverse selection that showed households that experienced high-cost health events reported higher willingness to pay for insurance. A number of other factors were also examined, including history with insurance, income, education, family composition, gender, and health history. They found that insureds reported slightly higher willingness to pay than those who were uninsured, illustrating a positive impact of insurance education. Second, the nominal willingness to pay increased with income, but relative willingness to pay, as defined by percentage of household income, decreased with income. Moreover, respondents with higher education and males reported a higher willingness to pay for insurance, but in general family composition had no effect.

In contrast, other scholars[42] have found evidence that adverse selection may not be a concern in certain micro health insurance programs. In studying community-based health insurance in rural Senegal, Jutting found that the illness ratio, defined as the number of cases of illness per household in the past six months divided by number of household members, was not a significant contributing factor in explaining the enrollment rate. He included type of illness as a dummy variable to control for low-cost versus high-cost illnesses.

[41] Dror et.al (2007)
[42] Jutting (2004); Dror et al. (2005)

In another study, Dror et al.[43] collected field survey data for micro health insurance units in Philippines and also concluded that adverse selection was not a factor in causing increased utilization by insureds. Two pieces of evidence supported this conclusion. First, the morbidity of the insured group and the noninsured control group, as assessed by the number of illnesses reported in the past three months, did not differ. Second, within maternity-related claims, the number of pregnancies in the past five years was used as a proxy of the need for maternity services. The result was a determination that the rates were even slightly higher among the uninsured group, which was contrary to the expected outcome if adverse selection existed.

Nguyen and Knowles[44] used both objective and subjective measures for health status and found that in Vietnam neither measurement could explain health insurance coverage purchased in the health insurance market for school-age children and adolescent students. Therefore, they did not find evidence of adverse selection. In addition, they found that the level of wealth, education, and the community to which the household belonged were positively correlated with the probability of purchasing health insurance for their school-age children. Moreover, a female head of household also increased the probability of purchasing health insurance; this was consistent with the fact that females were in general more risk averse. Finally, the structure of the family also mattered in that households prioritized young children, male children, and those with more schooling in their purchase decisions.

Compared to previous literature, this study focused on a micro health insurance program in Pakistan and conducted an in-depth analysis on two related, yet distinct questions. First, I examined the trend of sustainability by comparing the claim experience between newly enrolled

---

[43] Dror et al. (2005)
[44] Nguyen and Knowles (2010)

households and renewed households. This approach helps to reveal the trends in risk portfolios in the renewal business and it further helps to shed light on trend of sustainability.

Second, I further detected for the existence of adverse selection by using a revised method from the traditional “positive correlation test”. Testing in the Pakistan program for adverse selection by using the traditional “positive correlation test” would not have worked. This is because the product in the program studied here had a simple design that gave the same coverage to everyone. Instead of using the “positive correlation test” between risk type and insurance coverage, I examined the positive correlation between risk type and policyholder’s renewal decision to give a piece of evidence for adverse selection. In addition, compared to the work by Jutting and Dror et al., I used more in-depth measurements to determine healthcare need. In this study I categorized different types of claims, including maternity related cases, into various risk types, and examined the extent of adverse selection in different types of claims.

Chapter 3

AKAM Program and Data

3.1 Introduction to the AKAM Program

The AKAM Microinsurance Initiative began in 2006 with support from the Bill and Melinda Gates Foundation. AKAM is owned by the Aga Khan Development Network (AKDN), and is one of the ten agencies for AKDN. The agencies of the AKDN are private, international, nondenominational development organizations. They work to improve the welfare and prospects of people in the developing world, particularly in Asia and Africa, without regard to faith, national origin, or gender. The underlying objectives of AKAM are to reduce poverty, diminish the vulnerability of poor populations, and alleviate economic and social exclusion. AKAM aims to improve the quality of life of people by helping them to improve their incomes, become self-reliant, and gain the skills needed to graduate into the mainstream financial markets.[45]

The AKAM Microinsurance Initiative started its pilot enrollment period for an annual micro health insurance policy in the Northern Area (NA)[46] of Pakistan in November 2007. NA, located in the mountainous part of Pakistan, is the country's northernmost political entity and has an estimated l.35 million people scattered across six districts. Making health insurance available to the region's poor for the first time is a milestone in the provision of social services to the area. In the three years after the program was launched, over 100,000 members had been enrolled, or about 7.5% of the local population.

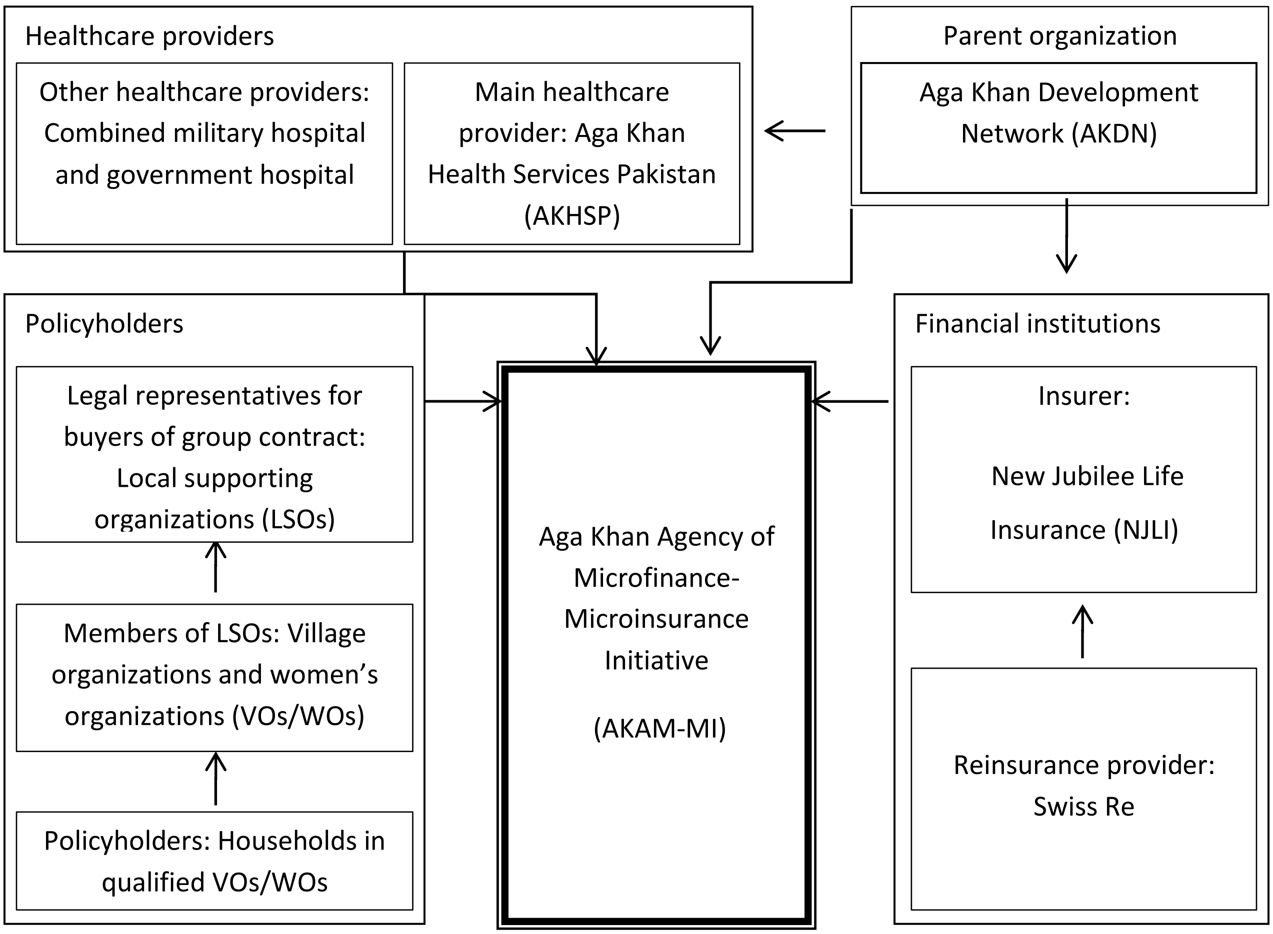
[45] www.akdn.org/about.asp

[46] The Northern Areas (NA) is now known as Gilgit-Baltistan (GB).

Figure 1 shows the flow chart for the operation and organizations associated with this micro health insurance program.

FIGURE 1

Flow Chart for AKAM Micro Health Insurance Program



The coverage and premiums are the same for every member, and the individual annual premium of 400 Pakistan rupees (PKR) (approximately $5.60) is paid up front.[47] There is no individual risk classification in underwriting. In exchange for this premium, the policy provides the following: annual hospitalization coverage (the core of the product) up to 25,000 PKR

[47] The annual premium was 350 PKR in November 2007. It increased to 400 PKR in November 2008 and stayed the same in July 2009 and November 2009. It increased again, to 450 PKR, for new insureds in two LSOs (ZADO and Danyore) in July 2010, but stayed the same for all the other insureds. In the data analysis, those households affected by the price increase were excluded from the sample, resulting in a loss of 193 observations.

(approximately $400),[48] life insurance of 25,000 PKR (approximately $400) on the head of the family, and one outpatient voucher valid for a one-time physician visit.[49] As an incentive to policyholders to renew their policies, the hospitalization coverage and the life insurance coverage increase to 30,000 PKR for renewed insureds for the same premium. The average annual income per person in the NA was around 50,000 PKR in 2009,[50] so the insurance coverage is equivalent to about half of a person's annual income.

AKAM partnered with local supporting organizations (LSOs) to distribute the product to voluntary groups. There were two reasons for selling group policies. First, the LSOs had existing networks for disseminating information; therefore, they could greatly reduce their underwriting and distribution costs. Second, reflecting lessons learned in developed markets, the sale of group policies helps to address concerns of adverse selection because these groups are not formed solely to purchase health insurance. The major types of groups are village organizations (VOs) and women's organizations (WOs), and their members could purchase health insurance through the VOs/WOs if at least 50% of the households in the organization agreed, in the test survey, to purchase the product if it were available. In addition, the entire household is required to enroll in the program in order to alleviate adverse selection.

AKAM relies on an existing network of VOs and WOs for product distribution. By the end of 2005, the NA had over 4,000 VOs and WOs that represented over 78% of the total households.[51] VOs/WOs are grass-root community organizations that were first developed to improve the capability of households to undertake village development initiatives. For example,

[48] 25000 PKR converted to $400 in November 2007, $312 in November 2008, and $293 in February 2011.
[49] AKAM pays the hospital 50 PKR for each voucher used. The outpatient voucher could be sold or transferred to others.
[50] http://www.unicef.org
[51] The Aga Khan Rural Support Program: an assessment of the institutional development of village and women's organizations, the AKRSP's Institutional Development Survey 2006.

projects have included obtaining funding and organizing the villagers to work on infrastructure projects such as minor irrigation works, flood protection, erosion control, and linking of roads. In addition, VOs/WOs were also involved in organizing informal community-based micro loans among their members before AKAM started its formal microfinance service in the area. More recently, these organizations have become key players in providing supportive networks to enlarge communities' assets and harness individuals' skills to generate sustainable forms of income.

LSOs are large organizations working for the member organizations in the area and may consist of 50 or more villages. They are nonprofit organizations set up in a joint effort with the Aga Khan Rural Support Program (AKRSP) and the local population and serve as registered legal entities under Pakistani law. LSOs are the vehicles for VOs and WOs, which are not registered organizations, to enter into legal agreements as a group or subgroup. Gradually, LSOs have obtained project funding and broadened their services.

AKAM initiated policies for New Jubilee Life Insurance Company (NJLI), which is a commercial insurer based in Karachi, Pakistan, and also owned by AKDN.[52] The LSOs enter into an agreement with AKAM, which has been appointed by NJLI to represent it in all matters pertaining to the health microinsurance program. The LSOs contract with NJLI on behalf of households living in VOs/WOs.

AKAM chose the NA as the first area in which to provide health insurance because of the high profile AKDN has there. The main health service provider, Aga Khan Health Services Pakistan (AKHSP), which is a part of AKDN, has operated in the NA for over 30 years. It has

[52] In regard to the reinsurance arrangements, AKAM arranged a stop-loss contract with Swiss Re.

three hospitals and 25 primary care facilities in the NA, and over 90% of the claims from the AKAM micro health insurance program are handled within AKHSP systems. In addition, the Combined Military Hospital (CMH) and a government hospital are located in the NA; these two facilities handle less than 10% of the claims, with most of these claims associated with emergency services.

3.2 Data Description

3.2.1 Data Structure and Summary

The empirical application uses a dataset composed of household level information, including some basic demographics on the age and gender of the head of the household, members' ages and gender, and household size. The VO and LSO to which a household belongs are also recorded. The dataset also has some policy-level information such as policy limit, renewal status, and enrollment date. Moreover, it contains detailed information on claims made during the policy periods. This information includes diagnosis, any pre-existing condition, the name of the hospital used, date of admission, length of stay, and total bill.

There have been six enrollment periods since the program was launched in November 2007. Each policy lasts for one year. In the beginning, the enrollment window was open only for a month, in November, to help alleviate adverse selection by not allowing people to buy insurance policies immediately after they learned surgery was needed. In addition, the one-month enrollment period reduced administrative and distribution costs. In addition to the four November waves of enrollment from 2007 through 2010, a July enrollment window was opened in 2009 because of the increasing demand for health insurance and greater knowledge of the income cycle for local households.[53] The policy was made available to households in the same VOs/WOs that met the minimum enrollment requirement in the previous November. The program also expanded into a few new VOs/WOs in the same area during the July enrollment periods.

The first enrollment period in November 2007 was a pilot program; therefore, detailed information on the household level was not collected at the time. Because of this lack of data, I omitted the first period from the regression data analysis in chapters 4 and 5, but included aggregated information in Figure 1 to provide a reference for the development of a loss ratio over years. For the fifth enrollment period that began July 2010, claim data was available only through November 2010. The sixth enrollment period, which started in November 2010, is excluded from the analysis because of lack of access to the data.

In summary, the data analysis is based on data from four available enrollment periods: November 2008, July 2009, November 2009, and July 2010. The data from these four periods total 15,962 household-year observations and 62,998 member-year observations.

Because the two key variables to examine are claim experience (in terms of frequency, severity, and total amount) and renewal status, detailed information on those variables across different enrollment periods is summarized in Table 1.

[53] Originally November was chosen as the month to open the enrollment window because after the fall harvest most of the farmers had cash to pay the premium. Then it was learned that another group of local residents received most of their income in the summer from tourism, so a July enrollment month was added to accommodate these residents.

TABLE 1

Summary of Household-level Claim Information and Loss Ratios for all Enrollment Periods[54]

Variable	2007 Nov	2008 Nov	2009 Jul	2009 Nov	2010 Jul*	Overall
Households enrolled	1,715	5,272	2,197	5,784	2,709	17,677
Members enrolled	6,044	19,463	9,087	22,940	11,508	69,042
Number of claims	768	3,370	1,905	3,474	891	10,408
Total claims (in PKR)	2,820,595	14,996,230	8,905,712	17,622,975	3,742,842	48,088,354
Total premiums (in PKR)	2,417,600	7,785,200	3,634,800	9,176,000	4,603,200	27,616,800
Loss ratio	1.17	1.93	2.45	1.90	n.a.	1.93

Table 1 shows that the claim experience and loss ratio fluctuated across different enrollment periods. For the three November waves of enrollment, the claim experience and its corresponding loss ratio increased in 2008 and then stayed about the same in 2009. The overall loss ratio for the first four enrollment periods was 1.93.

The AKAM program expanded over time into new areas and cooperated with new LSOs to deliver the product, which in general explains the increase in the number of household enrolled for the November cohort. The first five enrollment periods have a total of 14 LSOs, and the LSOs overlap across the November and July enrollment cohorts. There is a different pattern in the number of household enrolled in November versus the July cohorts. Except for the first November enrollment period, in general the November cohort enrollees outnumber the July cohort enrollees. This is because of two reasons. First, the November enrollment window was opened first in 2007, and the July enrollment window was not added until 2009. Second, the

[54] For the 2010 July enrollment period, there is only partial year (5 months) claim data (from July 2010 until November 2010). The number of claims and total claim amounts are actual five-month's data without adjustment, thus the loss ratio for the 2010 July period is not available. Total claim amounts are related to the policies initiated at each enrollment period.

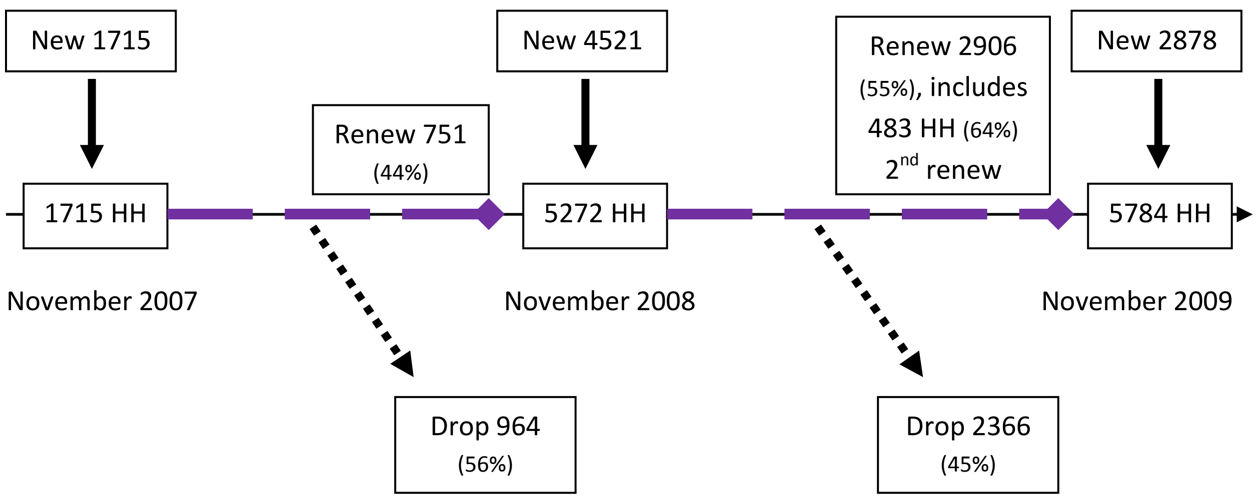
November cohort expanded into a larger area and contracted with more LSOs than the July cohort. For detailed information regarding LSOs entry information, please refer to Appendix 3.

Households' decisions to opt in/out and renew a policy are the key variable of interest. Figure 2 describes the timeline of the household enrollment structure for both the November and July cohorts.

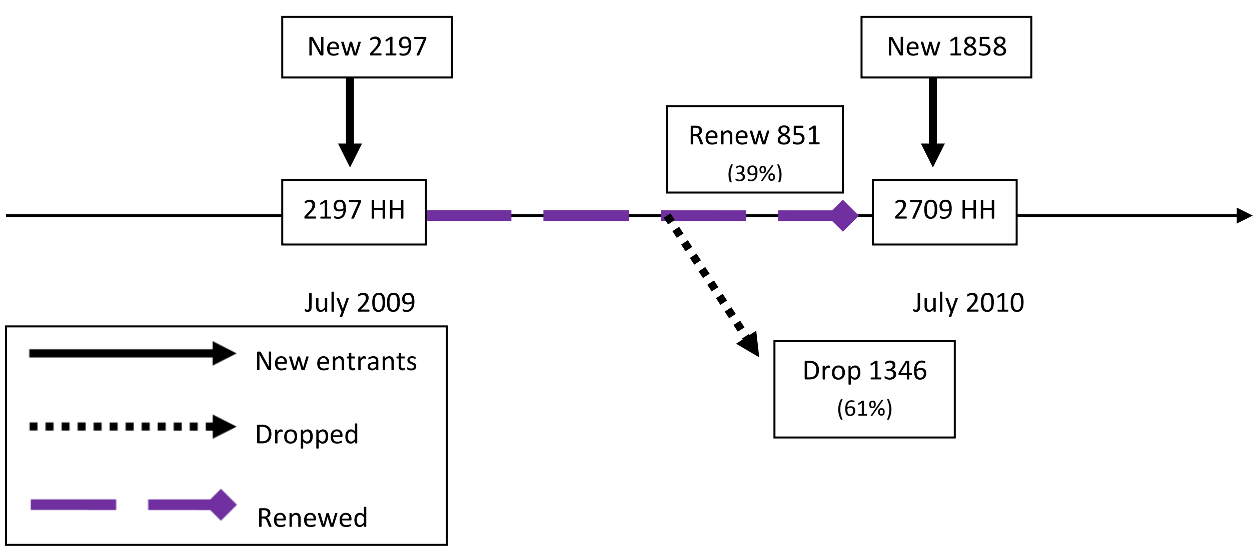
Figure 2

Timeline of Household Enrollment Structure

November cohorts' enrollment structure



July cohorts' enrollment structure



All households were new in the first enrollment waves in both July and November, but enrollment in the other three periods was a combination of new households and households that were renewing from previous periods. The renewal rate, defined as the percentage of households that chose to renew their policies at the end of the policy term, shows that the program has retained on average about half of the enrolled households (renewal rates are 44%, 55%, and 39%, respectively). Of those that renewed their policies once in November 2008, 64% (483 of 751 households) renewed for a second time in November 2009. However, the program was expanding rapidly from period to period, so in general, new households still were weighted more heavily than renewing households.

Table 2 describes all variables and summarizes statistics by enrollment period. The original household year observations totaled 16,225; 70 observations were excluded because they lacked birth year information for the head of household. In July 2010, the premium for new households enrolled in Danyore and ZADO was increased to 450 PKR from 400 PKR. For consistency and to exclude the impact of the price increase on a household's decision to enroll, an additional 193 households were excluded from the sample, resulting in a total sample size of 15,962.

TABLE 2

Variable Descriptions and Summary Statistics

Panel A: Demographics					
Variable	Full sample	2008 Nov	2009 July	2009 Nov	2010 July
HH size	3.99	3.73	4.21	4.01	4.28
	(1.75)	(1.67)	(1.70)	(1.76)	(1.86)
Head age	45.20	45.07	42.79	46.48	44.68
	(16.02)	(14.22)	(13.96)	(17.41)	(17.48)
Head male	0.73	0.74	0.74	0.72	0.75
Fraction child (age 0-5)	0.13	0.12	0.14	0.13	0.14
	(0.18)	(0.18)	(0.18)	(0.18)	(0.18)
Low-risk chronic	0.03	0.04	0.04	0.03	0.02
Medium-risk chronic	0.05	0.06	0.05	0.04	0.02
high-risk chronic	0.04	0.04	0.06	0.04	0.03
Pre-existing condition	0.08	0.09	0.11	0.08	0.04
Fraction teenager (age 6-16)	0.18	0.19	0.16	0.18	0.19
	(0.22)	(0.23)	(0.21)	(0.22)	(0.22)
Fraction adult (age17-59)	0.58	0.58	0.60	0.57	0.57
	(0.26)	(0.26)	(0.25)	(0.26)	(0.25)
Fraction elderly (age60-99)	0.11	0.12	0.10	0.11	0.10
	(0.21)	(0.22)	(0.17)	(0.21)	(0.20)
Fraction female	0.55	0.56	0.55	0.56	0.54
	(0.20)	(0.21)	(0.19)	(0.21)	(0.19)
Panel B: Claim					
Variable	Full sample	2008 Nov	2009 July	2009 Nov	2010 July
Claim frequency	0.60	0.64	0.87	0.60	0.33
	(1.01)	(1.02)	(1.23)	(0.99)	(0.71)
Claim severity	1861.06	1894.69	2483.10	2007.61	978.22
	(4050.13)	(3968.44)	(4374.89)	(4329.52)	(3055.58)
Total claim amount	2,834.99	2,844.51	4,053.58	3,044.16	1,381.63
	(6150.31)	(5851.72)	(7118.47)	(6603.94)	(4319.61)
Loss ratio	2.01	1.93	2.45	1.90	n.a.
N (in household year)	15,962	5,272	2,197	5,784	2,709

Note: Mean is shown on top, and the standard deviation for each variable is shown in parentheses.

3.2.2 Claim Data Risk Type Categorization

The dataset contains a total of 9,683 claim data categorized in 213 different types of diagnoses. In order to test for adverse selection in Chapter 5, I used a type of risk indicator to serve as a proxy for overall healthcare needs in the next year.

I went through four steps to categorize households into different types of risk according to individual members' claim histories. First, I divided all 213 different types of diagnoses into two categories, acute and chronic diseases.[55] Specifically, an acute disease is a short-term, curable sickness that enables patients with very limited ability to estimate their medical expenses in the next year, such as in cases of appendicitis and fractures. In contrast, chronic disease is a long-term, recurrent disease that would cause a policyholder to rationally expect some future medical cost in the next year. Some typical examples of chronic diseases include asthma, diabetes, cancer, and AIDS/HIV. The dataset contains 113 types of acute disease and 100 types of chronic disease.

Second, I further categorized all chronic diseases into three different levels of risk, because the treatments for different types of chronic disease could result in large variations. For example, the cost to treat asthma versus AIDS/HIV would differ greatly; therefore, I would expect households to have varying levels of motivation to renew their policy. In particular, I ranked the average cost for treatment of each diagnosis from the highest to the lowest, then the top 33% of

[55] There is not a standard, well-recognized division between acute and chronic disease; therefore, in conducting this project, I consulted with a healthcare provider with experience in developing countries.

diagnosis type was categorized as high-risk chronic disease, the middle 33% as medium-risk chronic disease, and the bottom 33% as low-risk chronic disease.[56]

TABLE 3

Summary of Claim Risk Types

Diagnosis category	Risk type	Examples
Miscellaneous	0	Claim without diagnostic information, claim with unknown diagnosis, chronic disease claim with fewer than three cases
Acute disease	1	Appendicitis, fracture
Low-risk chronic disease	2	Asthma
Medium-risk chronic disease	3	COPD, sepsis
High-risk chronic disease	4	Renal failure

TABLE 4

Summary Statistics of Total Bill by Household Risk Types

Category	Risk type	Observation	Average total bill	Std. Dev.	Min	Max
0	Miscellaneous	10355	202	1,735	0	30,000
1	Acute	3,839	5,356	6,362	119	59,629
2	Low-risk chronic	537	6,760	6,617	205	47,501
3	Medium-risk chronic	735	9,788	8,808	138	56,549
4	High-risk chronic	689	17,469	8,835	417	58,836
Overall		**16,155**	**2,817**	**6,136**	**0**	**59,629**

Third, the claims with no diagnostic information or with an unknown diagnosis were categorized as risk type 0,[57] and the acute diseases, low-risk, medium-risk, and high-risk chronic

[56] Of 100 types of chronic disease, 31 were dropped because of less than three occurrences in claim frequency in the dataset, resulting in an inability to calculate a representative average cost for treatment. Their elimination resulted in the loss of 45 claim records, i.e. 0.5% of the total number of claims. The remaining 69 types of chronic disease, all of which occurred at least three times, were further divided into low-, medium-, and high-risk types, with each category having 23 types of chronic diseases.

diseases were designated, respectively, as risk types 1, 2, 3, and 4. Table 3 illustrates the summary of risk types. I used the average cost of treatment for chronic disease as a proxy for their likely recurrent future cost. The type of risk increases as this estimated future treatment cost increases.

The last step used the individual claim records of members of a household to assess the risk posed by the household. The risk type for each household is the highest category of all household members' individual claim types. Table 4 summarizes the statistics of the total bill by household risk type. The average of total bills across different risk categories roughly agrees with our expectation that higher risk categories should incur higher average total bills. In the regression analysis in Chapter 5, I excluded all households in risk type 0 from the sample size and used households in risk type 1 as a base category for comparison with the other three types of households. Please refer to Appendix 1 for a complete list of the categories of risk types, claim frequency, and average treatment cost.

[57] The 31 types of chronic disease that were dropped were also assigned risk type 0.

Chapter 4

Development and Sustainability in the Micro Health Insurance Market: Evidence from Pakistan

4.1 Introduction

The insurance industry in Pakistan is still in its early stages and has a slow growth rate. In 2008, insurance penetration was only 0.7% of the GDP, and the insurance density was merely $6.50 for the entire industry, let alone its infant microinsurance section.[58] Many microinsurance programs had high loss ratios and lapse rates in their early years, and these losses and lapses gave rise to concern among donors and management teams about the issue of sustainability.

From an actuarial perspective, it is crucial to understand the claim trends in the renewal business as a prerequisite for a thorough analysis of sustainability; moreover, from a risk management perspective, it is important to analyze whether there are adverse selection issues that lead toward a "death spiral" of the market.

In this chapter, I examined the viability of micro health insurance by looking into households' renewal decisions. Specifically, I analyzed the trend of sustainability by measuring the development of claim experiences in renewal policies. I used a two-part model to compare the microinsurer's new book of business with its renewed book of business and to test whether

[58] Securities and Exchange Commission of Pakistan Annual Report 2009, p 109

the risk pool had deteriorated over time. The status of the risk pool sheds light on the sustainability of the microinsurer's business.

From a theoretical standpoint, the development of renewal policies could result in either increasing or decreasing sustainability. On the one hand, there are forces leading to decreased sustainability via the renewal business in the sense that people with higher risks might be more likely to stay with an insurer. One classic version of the adverse selection story here would be if policyholders had private information about their risk types and suffered unrelated income shocks that affected whether they could afford insurance. In that case, I would expect that those who know they are at especially high risk would be less likely to cancel insurance because of an income shock; retention of such policyholders would result in a deteriorating book of business in terms of renewed policies. Another possibility motivating people at higher risk to stay with a program is that people might learn about their risk type over time – for instance, learning about a pregnancy. If people act on that new information, I would expect adverse selection because those aware of their higher risk would choose to renew but low risks may not.

On the other hand, other forces that affect the renewal decision may be unrelated to risk type and could even result in increasing sustainability. First, more financially savvy people might be more likely to renew; if they are also healthier, then their renewals improve the risk portfolio in terms of sustainability. Second, people who are more risk averse are more likely to renew; and these policyholders may be willing to take more preventive care and act more cautiously, which may make them healthier, and, therefore, these factors could also improve sustainability. Last but not least, in new markets like Pakistan, people start to learn the value of insurance and how it protects their financial stability and flexibility. People who experience a major medical incident that might be unlikely to happen again (such as appendectomies and cholecystectomies) may still want to renew because they better understand and appreciate the value of insurance. These people also could well be healthier because of the treatment received; therefore, their renewal could improve the risk portfolio. Ultimately, which direction to pursue to achieve sustainability in this type of microinsurance market remains an open empirical question.

I analyzed how claim rates evolve as households renew their policies and found that households that had larger claims during the initial policy period were slightly more likely to renew their policy for the next period. Although that pattern is superficially consistent with adverse selection and decreasing sustainability, I found in delving deeper that when compared with households buying insurance for the first time, renewing households had significantly lower claim frequency and total claim amounts. Taken together, these results suggest there are forces affecting the demand for insurance within renewing households that may lead to an improved risk portfolio. This could also be explained by the prevalence of pre-existing conditions among newly enrolled households who have access to affordable healthcare for the first time. In addition, some selection issues may also contribute to the high claim frequency for newly enrolled households.

4.2 Analytical and Statistical Modeling

To evaluate the evolution of claim experience as policyholders renew their policies, on the one hand, classic models of adverse selection that predict that high-risk individuals are more likely to purchase and renew coverage than low-risk individuals, thus leading to a deteriorating risk portfolio and to decreasing sustainability. On the other hand, a number of empirical studies in developed insurance markets document “advantageous selection” in which low-risk individuals are more likely to purchase and retain the most insurance, thus leading to improving portfolios and sustainability.[59] In the face of these conflicting perspectives, theory and empirical evidence suggest that either pattern of sustainability is possible—making it all the more important to test for the nature of selection in developing health insurance markets.

The main focus of this study is on the relationship between a policyholder’s renewal decisions and the household’s healthcare claim expenditures. In the health literature, the modeling of healthcare expenditures could take either of two routes. The first involves taking expenditures as one part of a model and applying censored and truncated regression models, such as Tobit, to address the fact that healthcare expenditures cannot be negative.[60] The other approach is to treat expenditures as consisting of two parts, namely frequency and severity, so that these two important features can be modeled separately.[61] A frequency model takes into consideration a large number of zeros in the claim data and models the probability of filing a claim with either a probit or logit regression. The severity model takes into consideration that the expenditure is usually skewed in the right tail and models the nonzero expenditure with either ordinary least square (OLS) on the log of the expenditures or with general linear models (GLM) estimators (usually Gamma with a log link function). In this paper, I apply both one- and two-part model approaches to the modeling of claim expenditures, and I used three models.

The first model uses binary logit to focus on how claim experience during the current period affects a household’s decision to renew the policy in the next period. It is assumed that the probability of renewing $p_{i,t+1}$ depends on the following factors: total claim amount filed from

---

[59] Bolhaar et al. (2008); Fang et al. (2008); Gao et al. (2009); Einav and Finkelstein (2011)
[60] Amemiya (1984)
[61] Gao (2007), Manning et al. (2004)

the current period $z_{i,t}$, whether the household was renewed at the beginning of current period $R_{i,t}$, household characteristics $H_{i,t}$, community characteristics $C_{i,t}$, and the error term u, which is a random variable.

Model 1 could be described as follows:

Model 1 (Probability of Renewing):

$$p_{i,t+1} = f\left(z_{i,t}, R_{i,t}, H_{i,t}, C_{i,t}, u\right) \quad (1)$$

Specifically, $R_{i,t}$ is a renewal dummy to indicate whether household i was new or renewed at the beginning of the period t. Household characteristics $H_{i,t}$ include family size, age and gender of the head of household, and the percentages of children, teenagers, elderly, and females in the family. Community characteristics $C_{i,t}$ are the LSO dummies that designate the LSO that the household belongs to. These dummies are included to account for the geographical differences that cause variations in the cost of seeking healthcare.

To estimate the probability of renewing, I used a binary logit model:

$$p_i{}^* = (\beta z_i + \emptyset R_i + \alpha H_i + \ \delta C_i + u_i) \quad (2)$$

$$p_i = 1 \; if \; p^* > 0, meaning \; the \; household \; i \; will \; renew \; in \; the \; next \; period$$

$$p_i = 0 \; otherwise$$

The other two models focused on measuring the differences in claim experiences between newly insured households and renewed households. Model 2 used a two-part frequency severity model as a first approach to measuring these differences. In the second approach to this assessment, Model 3 relied on a Tobit model to examine the total claim amount.

Finding a positive relationship between policy renewal and claims would be consistent with a dominant role of adverse selection in the decision to retain insurance. A negative correlation between renewal and claims, however, would suggest other factors motivated the healthier insureds to retain their insurance coverage.

In summary, models 2 and 3 took the following forms:

Model 2 (Two-part model):

$$Claim\ frequency_{i,t} = f(\ R_{i,t}, H_{i,t}, C_{i,t}, E_{i,t}, u) \quad (3)$$

$$Claim\quad severity_{i,t} = f(\ R_{i,t}, H_{i,t}, C_{i,t}, E_{i,t}, u) \quad (4)$$

Model 3 (Total claim model):

$$Claim\ amount_{i,t} = f(\ R_{i,t}, H_{i,t}, C_{i,t}, E_{i,t}, u) \quad (5)$$

The primary measures of claim experience are claim frequency, claim severity, and total claim amount; the latter equals the product of claim frequency and severity. In the two-part model, I used claim frequency, defined as a dummy variable indicating that the household filed at least one claim in the policy year, as the dependent variable. I modeled it with logit regression in the first part. Then, conditioned on the household's having at least one claim, I used claim severity, defined as total claim amount for the household in a policy year, as the dependent variable. Then I used OLS to model the natural log transformation of claim severity.[62] As an alternative to the two-part model, I also included Model 3, using annual claim amounts as the dependent variable, as a sensitivity test. For this specification I used a Tobit model left censored at zero to take into account the presence of a large number of zeros in the dependent variable.

[62] In the two-part model, the primary concern is not the full impact of renewal status on the total claim amount, but the component effect of renewal status on frequency and severity separately.

In both models 2 and 3, I began by using only renewal status $R_{i,t}$ as the independent variable in specification [1], and then I also included additional control variables to measure other policyholder characteristics that might affect claim experience throughout models 2 and 3 in specification [2]. The household characteristics $H_{i,t}$ were the log of family size, age of the head of household, gender of the head of the household, and family structure (percentages of members that were children, teenagers, elderly, and female). The community variables $C_{i,t}$ are dummy variables indicating the LSO that the insured household belonged to. The enrollment period dummy variables $E_{i,t}$ are also included to control for unobserved time effects.

4.3 Empirical Results

The first section of my presentation of results shows a univariate comparison of claim frequency, severity and loss ratio by renewal status. The second section gives an analysis of loss ratio and the last section presents the regression results from the models.

4.3.1 Univariant Comparison Analysis

Table 5 summarizes the mean in terms of renewal status. It shows that renewed households had a better claim experience than newly enrolled households. These renewed households had a lower probability of filing a claim, a lower average claim severity, and better loss ratios. The average claim frequency here is defined as the proportion of households that filed at least one claim. This frequency decreased from 39% among newly enrolled households to 34% and 36%, respectively, among first-time and second-time renewed households. A similar trend appears in claim severity. The average claim severity is lower among first-time and second-time renewed households (2,933 and 3,109 respectively) compared with 3,247 in the case of newly enrolled households. However, it should be noted that the first-time renewed households seem to show better claim experience than the second-time renewed households. But we also observed that renewed households on average had more family members than the new households, with more males as the heads of the family and higher percentages of teenagers and elderly. This last observation raises the question of whether the claim trend would be the same if all these other factors were controlled for.

TABLE 5

Mean by Renewal Status

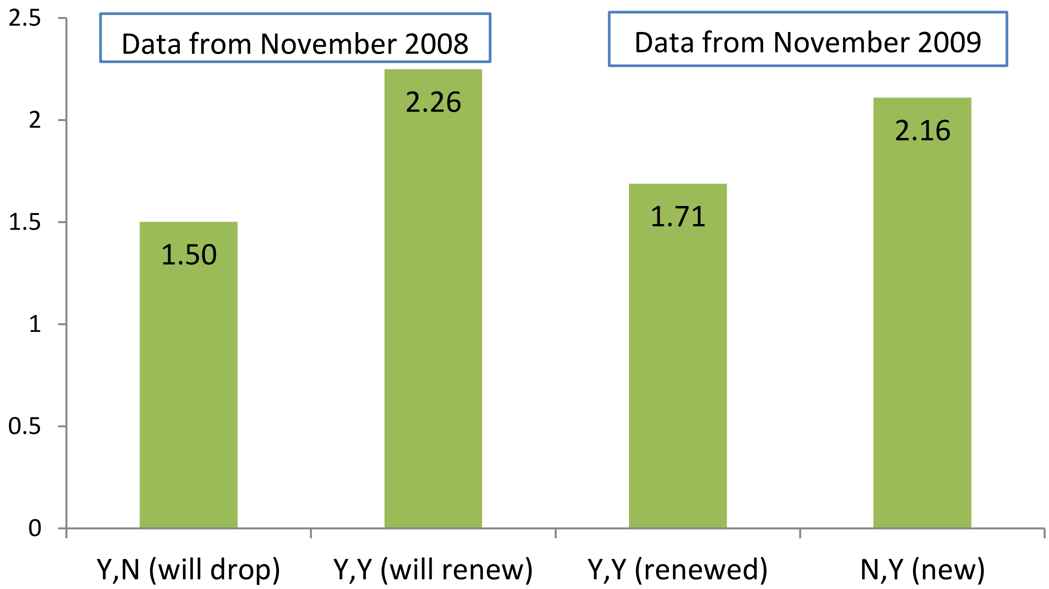
Variable	New household	1st time renewed household	2nd time renewed household
Claim frequency	0.39	0.34	0.36
Claim severity	3,247	2,933	3,109
Loss ratio	2.13	1.71	1.79
Family size	3.81	4.29	4.35
Age of head of household	45.07	45.26	47.92
Head male	0.73	0.75	0.80
Fraction child	0.13	0.14	0.12
Fraction teenager	0.18	0.19	0.24
Fraction adult	0.59	0.54	0.52
Fraction elderly	0.10	0.12	0.12
Fraction female	0.56	0.55	0.54
N	11,454	4,025	483

4.3.2 Loss Ratio Comparison Analysis

To summarize the trend in terms of sustainability and to compare directly the dynamics of claim experience between newly enrolled and renewed households, I took the November cohorts and divided them into four groups to compare their loss ratios, as shown in Figure 3. These results derived from November cohorts observations are shown in the graph below, and they foreshadow the results from regression analysis in the next section.

FIGURE 3

Loss Ratio Comparison for November Cohorts



In Figure 3, I divided all enrolled households in November 2008 and November 2009 into four groups according to their renewal decisions in November 2009. The two columns on the left represent data from November 2008 and the two columns on the right are from November 2009. Then within each period I divided the insured households according to their renewal status in November 2009. In particular, the first column shows the loss ratio for those households enrolled in November 2008 that will drop their policies in November 2009; the second column shows loss ratios in November 2008 for those households that will renew in November 2009. It is clear that those who chose to renew for the next period have a higher loss ratio than those who chose not to renew (2.26 compared to 1.50). This comparison suggests a likely adverse selection scenario.

The third column represents the loss ratio in November 2009 for those households that renewed in November 2008. This means these households are the same as those in column 2, but with a loss ratio updated in November 2009. It is obvious that the loss ratio of this same group of households improved dramatically in a year (a decrease from 2.26 to 1.71), making their loss ratio quite comparable to the first group. Moreover, their loss ratio is much lower than the loss ratio of newly enrolled households (1.71 compared with 2.16 in column 4).

Similar patterns occur in comparisons using either total claim amounts or claim frequency. In November 2008, the group that will renew had a higher claim frequency, a higher total bill, and a very high loss ratio in comparison with the group that will not renew. However, a year later in November 2009, this same group of people actually had a much lower claim frequency, a lower total bill, and much improved loss ratio than in the previous period. These November 2009 figures for renewals are actually much lower than the same figures for the newly enrolled households. These patterns illustrate some of the key findings of regression models that are discussed in the next part.

4.3.3 Regression Results

The regression results for the three models are shown in tables 6-8, respectively.

TABLE 6

Regression Results for Model 1 (Probability of Renewing)

Dependent variable=1 if HH is going to renew for the next period				
	OLS		logit	
Variable	[1]	[2]	[1]	[2]
Total bill (in thousands)	0.01*** (0.001)	X	0.01*** (0.001)	X
Had claim	X	0.12*** (0.001)	X	0.13*** (0.001)
Previously renewed	0.18*** (0.001)	0.18*** (0.001)	0.18*** (0.001)	0.18*** (0.001)
Log (family size)	0.10*** (0.001)	0.10*** (0.001)	0.11*** (0.001)	0.11*** (0.001)
Head age	-0.0004 (0.436)	-0.0001 (0.838)	-0.0004 (0.439)	-0.0001 (0.827)
Head male	-0.011 (0.507)	-0.007 (0.651)	-0.011 (0.506)	-0.008 (0.659)
Fraction child	0.08** (0.034)	0.06 (0.119)	0.09** (0.037)	0.06 (0.131)
Fraction teenager	-0.09*** (0.007)	-0.08** (0.018)	-0.10*** (0.006)	-0.09** (0.016)
Fraction elderly	0.11*** (0.001)	0.12*** (0.001)	0.12*** (0.001)	0.13*** (0.001)
Fraction female	-0.08*** (0.010)	-0.09*** (0.004)	-0.09*** (0.010)	-0.10*** (0.004)
July 2009	-0.15*** (0.001)	-0.15*** (0.001)	-0.16*** (0.001)	-0.16*** (0.001)
LSO fixed effect	X	X	X	X
R square	0.545	0.547		
AIC	10,221	10,195	9,729	9,704
N	7,469	7,469	7,469	7,469

Note: The dependent variable is a dummy variable indicating whether a household renewed in the next period, and I ran both an OLS and a logit model (marginal effects reported) using two specifications: [1] with total bills filed during a current policy period and other control variables, and [2] with a dummy variable indicating at least one claim was filed during a current policy period and other control variables. P values are shown in parentheses.

*significant at 10%; **significant at 5%; ***significant at 1%

In Table 6, using the enrollment data from November 2008 and July 2009 together, I found that both the total bill during the policy year and an indicator that the household had at least one claim were positively correlated to the renewal decision for the next period. When all other factors were held constant, every 1,000 PKR increase in claims in the current period increased the probability of renewing by 1%. Moreover, when other factors were held constant, those households that had at least one claim during their policy period increased their propensity to renew by 13%.

Households with larger claim amounts in one period are more likely to renew their policy in the next. This could happen for a number of reasons. First, traditional adverse selection scenarios may be at work in which high-risk households have private information and thus choose to renew their policies. Another possibility is that households that are not consistently at high risk happened to experience claims in a current period. The utilization of insurance makes them value the policy, therefore, leading to decisions to renew. To delineate between these two possibilities, I further examined the relationships in Model 2, which breaks claim amounts into frequency and severity to check their relationship with renewal status, and in Model 3, which uses the claim amount itself. The results are reported, respectively, in tables 7 and 8.

In Table 6, I also show that a renewal decision for the following period is positively affected by a renewal decision made in the previous period. Those households that renewed at the beginning of the previous period were 18% more likely to renew for the next period when other factors remained constant.

Family size is another important factor in the decision to renew. I used the log of the total member count in a household as a measurement of family size to account for the nonlinearity between the member count and decision making. I found that when other factors were held constant, every 10% increase in family size resulted, on average, in a 1.1% increase in the probability a policy would be renewed.

As for the impact of age composition, both the percentages of children and of elderly in the family had a consistently positive impact on renewal decisions. Every 10% increase in the percentage of children and elderly in a household, when other factors were held constant, increased the probability of policy renewal by 0.9% and 1.2%, respectively. The impact of the percentage of teenagers in households was negative. Each 10% increase in the percentage of teenagers in a household decreased the probability of renewal by 1%. Similarly, the percentage of females in a household had a consistently negative impact on renewal probability; each 10% increase in the percentage of females decreased the probability of renewal by 0.9%, all else being equal.

TABLE 7

Regression Results for Model 2 (Two-part Model)

	Frequency model		Severity model			
	Dependent Variable=1 if filed any claim		ln (claim amount)		claim amount	
	logit		OLS		GLM-gamma	
Variable	[1]	[2]	[1]	[2]	[1]	[2]
1st renewal	-0.05***	-0.037***	0.05	-0.05	0.04	-0.07**
	(0.001)	(0.001)	(0.124)	(0.132)	(0.314)	(0.048)
2nd renewal	-0.04	-0.043*	0.12	0.01	0.05	-0.03
	(0.109)	(0.062)	(0.115)	(0.888)	(0.514)	(0.735)
Log (family size)		0.16***		0.35***		0.33***
		(0.001)		(0.001)		(0.001)
Head age		-0.003***		0.002**		0.002**
		(0.0010)		(0.021)		(0.050)
Head male		-0.003		-0.05		0.03
		(0.767)		(0.205)		(0.399)
Fraction child		0.17***		-0.49***		-0.53***
		(0.001)		(0.001)		(0.001)
Fraction teenager		-0.31***		-0.41***		-0.31***
		(0.001)		(0.001)		(0.001)
Fraction elderly		0.03		0.40***		0.36***
		(0.187)		(0.001)		(0.001)
Fraction female		0.08***		-0.17**		-0.12
		(0.001)		(0.030)		(0.120)
July 2009		0.01		0.050		0.09**
		(0.585)		(0.220)		(0.031)
November 2009		-0.02*		0.07**		0.11**
		(0.100)		(0.028)		(0.003)
July 2010		-0.24***		0.56***		0.71***
		(0.001)		(0.001)		(0.001)
LSO fixed effect		X		X		X
R square			0.0007	0.9859		
AIC	21,139	19,693	17,826	17,331	121,036	120,515
N	15,962	15,962	6,032	6,032	6,032	6,032

Note: For the frequency model, the dependent variable is a dummy variable indicating at least one claim was filed during the policy period, and I ran a logit model (marginal effects reported) using two specifications: [1] with only renewal status as an independent variable, and [2] with other control variables. In conditions under which the claim frequency is greater than zero, I ran the severity model with ln (claim amount) as the dependent variable with OLS, as well as with GLM with a Gamma distribution and a log link function. P values are shown in parentheses.
*significant at 10%; **significant at 5%; ***significant at 1%

Table 7 demonstrates the regression results for the two-part model. The analysis of claim frequency shows that renewed households were less likely to file claims than new customers. That pattern is confirmed in both specifications in Table 7. According to a logit model, a first-time household renewal, all else being equal, decreased the probability of filing any claims by 3.7%, and a second-time renewal decreased this probability by 4.3%.

In addition, I found that larger families were more likely to file claims, which is intuitive. Each 10% increase in family size, all else being equal, led to a 1.6% increase in the probability of filing claims. Similarly, families with higher percentages of children and females had more claims. Among these families, all else being equal, each 10% increase, on average, affected the probability of filing claims by 1.7% in the case of children and 0.8% in the case of females. And each 10% increase in percentage of teenagers will decrease the probability of filing a claim by 3.1%, all else being equal.

I ran the severity model with a sample of all household who filed at least one claim during the policy year. Under the first specification, the dependent variable is the natural log of the total claim amount during the policy year. Using the OLS model, I found that the claims of households renewing for the first time were slightly less severe than those of newly enrolled households. However, this 5% difference between the two groups was not statistically significant when all control variables were included. Under the second specification, I ran a GLM with a Gamma distribution and a log link function using the total claim amount as the dependent variable. I found that compared to newly enrolled households, the status of being a first-time renewal household would decrease the severity of a claim by 7%, all else being equal, and the result is significant. And being a second-time renewal household would decrease claim severity by 3%, however, the result was not significant. Overall, the impact of a reduction on the claim experience is more significant for a first-time renewal household than it is for a second-time renewal household, and the significance level in the severity model was lower than in the frequency model.

The results of the two-part model are consistent with the health literature on selection. The existence of selection is expected in frequency analysis because households make their own decisions to seek medical services. However, selection is not necessarily expected in severity analysis because severity is also driven by the decisions of physicians, a situation exogenous to households' decisions.

As a sensitivity test for Model 2, I also ran another model for the total claim amount. The results for both the OLS and the Tobit model are presented in Table 8.

The results are by and large consistent with the two-part model, showing that both first-time and second-time renewals had significantly lower total claim amounts than those of newly enrolled households. According to the OLS model, compared with a newly enrolled household, a first-time household renewal, all else being equal, decreased the size of total claims by 27%, and a second-time renewal decreased the size of total claims by 32% (or a 5% decrease compared with a first-time renewed household).

The coefficient estimates for the other control variables are also consistent with the results in the previous models. Family size had a significant positive impact on total claims. The age of the head of household negatively affected the total claim amount. As for age composition, having a higher percentage of teenage family members decreased the total claim amount, and having more children and elderly in the household increased the total claim amount significantly. A higher percentage of female family members also increased the total claim amount.

TABLE 8

Regression Results for Model 3 (Total Loss Amount)

	Dependent variable: ln (total claim amount+1)		Dependent variable: total claim amount	
	OLS		Tobit	
Variable	[1]	[2]	[1]	[2]
1st renewal	-0.41***	-0.31***	-1429.33***	-1472.40***
	(0.001)	(0.001)	(0.001)	(0.001)
2nd renewal	-0.26	-0.38*	-876.89	-1420.41*
	(0.182)	(0.064)	(0.273)	(0.093)
Log (family size)		1.37***		6312.36***
		(0.001)		(0.001)
Head age		-0.02***		-65.05***
		(0.001)		(0.001)
Head male		-0.05		312.26
		(0.541)		(0.389)
Fraction child		1.26***		2411.76***
		(0.001)		(0.003)
Fraction teenager		-2.46***		-10683.91***
		(0.001)		(0.001)
Fraction elderly		0.37**		2261.30***
		(0.042)		(0.005)
Fraction female		0.57***		1965.78***
		(0.001)		(0.008)
July 2009		0.14		436.53
		(0.198)		(0.269)
November 2009		-0.09		-40.93
		(0.292)		(0.904)
July 2010		-1.96***		-5126.87***
		(0.001)		(0.001)
LSO fixed effect		X		X
R square	0.0018	0.4260		
AIC	90,805	89,448	143,717	142,561
N	15,962	15,962	15,962	15,962

Note: The dependent variable is the natural log of the total claim amount that a household incurred during the policy year plus one. I ran both an OLS and a Tobit model with the same two specifications as in Table 7. P values are shown in parentheses.

*significant at 10%; **significant at 5%; ***significant at 1%

In conclusion, the combination of the regression results and the observations in Figure 3 would seem to show even at first glance that the households that chose to renew were high risks, but this would not necessarily be true. Over time these renewed households were shown to be no worse than newly enrolled households. In fact, they filed fewer claims and had a lower total claim amount than newly enrolled households. This longer-term outcome is more consistent with the explanation that people who had claims valued insurance more and that this value judgment motivated them to renew instead of their renewals reflecting decisions motivated by a higher propensity to incur a claim. This is also consistent with the fact that people in developing countries have pent-up healthcare needs that emerge when health insurance is first provided. Because of this pent-up demand, the loss ratio for newly enrolled households is high; however, as time goes by, the risk portfolio may improve as these households renew their policies. Therefore, if this trend persists, it would support increased sustainability for healthcare insurance rather than decreasing its long-term sustainability.

4.4 Conclusion

In response to the need for affordable and quality health insurance for the poor, micro health insurance programs have been established with financial aid from donors in the hope that these programs will eventually become sustainable. Little empirical evidence is available about the sustainability of these infant microinsurance health insurance programs despite the critical need for data analysis that would advance our understanding of what is needed for these undertakings to succeed. In an effort to meet this need, I used data from a micro health insurance program in Pakistan to examine the relationship between claim histories and policy renewal decisions as a way to observe the development of renewal policies. I found that although households that filed higher claim amounts were more likely to renew their policies, these same households subsequently had fewer claims and lower total claim amounts than newly enrolled households. The renewed households had a much improved claim experience as time passed; moreover, they had comparable, if not better, claim frequency, total claims, and loss ratios than newly enrolled households. Thus, the entire risk portfolio actually improved over time instead of deteriorating as predicted by a classic adverse selection process.

This evidence suggests that part of the problem of high loss ratios stems from an initial surge of claims during the first year of coverage. This surge points to a conclusion that microinsurers need to realize that offering insurance for the first time among populations that may never have had affordable healthcare unleashes pent-up demand for health services to address a range of pre-existing conditions. These findings shed light on the possibility, despite a lack of underwriting, of long-term development of sustainable microinsurance programs in developing countries.

Chapter 5

Adverse Selection in the Micro Health Insurance Market: Evidence from Pakistan

5.1 Introduction

Adverse selection is one of the biggest barriers to the sustainability and success of micro health insurance programs.[63] Compared with the effect of adverse selection in the developed health insurance market, the persistence of adverse selection in the developing market could hinder this emerging industry from further development. Adverse selection, in combination with mistrust of insurers and lack of familiarity with insurance products, seems to be the reason for low sign-up rates, high claim rates, and low renewal rates.[64] Therefore, it is crucial to understand the extent of the existence of adverse selection as well as the effectiveness of ways to alleviate its impact.

Although selection is an important issue associated with sustainability, testing for the nature of selection in microinsurance markets presents unique challenges. Most studies of selection issues in insurance must overcome the inability to observe outcomes for people who do not purchase insurance. In developed markets, researchers typically approach these issues by using tests for adverse selection that analyze differences in claim experience across insureds who purchase different amounts of insurance.[65] In contrast, however, in developing markets, as with

[63] Biener and Eling (2012); Brau et al. (2011)
[64] Bendig and Arun (2011)
[65] Chiappori and Salanie (2000)

the AKAM product studied here, customers are given only one coverage option. This is done to hold down administrative costs and keep the products simple for a population inexperienced in purchasing insurance. Nevertheless, the consequence of these two factors is that the standard "positive correlation test" between coverage level and risk is not feasible.

Although some literature exists on adverse selection in micro health insurance, it falls far short of presenting a conclusive and convincing statement.[66] To address this issue in this paper, I primarily used detailed claim data from the AKAM micro health insurance plan and analyzed the relationship between households' risk types and their policy renewal decisions. Because the AKAM program is voluntary with virtually no individual underwriting required, adverse selection could be a major concern; however, the program imposes four requirements that target alleviation of adverse selection:

First, the health insurance product is sold as a voluntarily purchased group policy to households in Village Organizations or Women's Organizations (VOs/WOs). Second, the VOs/WOs must meet a 50% minimum enrollment ratio test, meaning that to be eligible to buy the product, at least half of the households in either organization must agree to participate. Third, an entire household must enroll so that both the sick and healthy members of a household are included. Last but not least, fixed enrollment windows occur in each November (and July),[67] which means that households must plan ahead instead of enrolling right before care is needed.

With these requirements in place, it is interesting to observe and test their effectiveness by examining the extent of adverse selection in this market. Furthermore, based on data analysis, I make suggestions for microinsurer to alleviate adverse selection and improve sustainability.

[66] Bendig and Arun (2011) ; Dror et al. (2005); Ito and Kono (2010); Lammers and Warmerdam (2010)
[67] Households could enroll in November from 2007 until 2010; the July enrollment window was available from 2009 until 2010.

5.2 Analytical and Statistical Modeling

***Hypothesis 1*:** In an insurance market characterized by adverse selection, high-risk households are more likely to renew their policies than low-risk households.

To test for adverse selection, I used information about households' risk types. This was derived from households' claim history in the current enrollment period, and it served as an indicator of the likelihood of incurring future medical expenses. This indicator was then regressed on a household's decision to renew the policy for the next period. According to a modified "positive correlation test"[68] of adverse selection, if high-risk households are shown to be more likely to renew their policies, then a positive correlation exists between risk type and insurance coverage purchased, and this correlation serves as evidence of adverse selection.

The main focus of this study is on the relationship between a household's risk type in the current period and its policy renewal decision for the next period. This relationship was examined through two models.

The first regression model takes the following form in which f is a binary logit regression with the dependent variable being a household's renewal decision for the next period. It is assumed that the probability of renewing $p_{i,t+1}$ depends on a household's risk type $W_{i,t}$, a household's renewal status in the current period $R_{i,t}$, its claim frequency $X_{i,t}$, a pre-existing condition dummy $Y_{i,t}$, household characteristics $H_{i,t}$, community characteristics $C_{i,t}$, and an error term u, which is a random variable[69].

---

[68] Rothchild and Stiglitz(1976); Puelz and Snow (1994)

[69] In theory, the latent risk type variable should absorb information provided by claim frequency and pre-existing conditions; however, here Z is the empirical risk type derived from historical claim data. It shows whether, under

Therefore, Model 1 could be described as follows:

$p_{i,t+1} = f(W_{i,t}, R_{i,t}, X_{i,t}, Y_{i,t}, H_{i,t}, C_{i,t}, u)$ (6)

Specifically, $W_{i,t}$ consists of three dummy variables that indicate whether the household belongs to a low-risk, medium-risk, or high-risk category. The omitted base category for comparison is those household that file only acute disease claims. $R_{i,t}$ is a renewal dummy to indicate whether the household was new or renewed at the beginning of the period. $X_{i,t}$ shows claim frequency, which is the total number of claims filed during the current period. $Y_{i,t}$ indicates whether there is a claim associated with a pre-existing condition. As noted earlier, household characteristics $H_{i,t}$ represents the makeup of the household, including family size, age and gender of the head of household, and the percentages of children, teenagers, elderly, and females in the family, while community characteristics $C_{i,t}$ are the LSO dummies that identify the local supporting organization that the household belongs to. These dummy variables are included to account for the geographical differences that cause variations in the cost of seeking healthcare.

In the second model I used logit regression again with the same control variables, but I restricted the sample to households having at least one GYN/OB claim. The purpose of Model 2 was to further investigate potential adverse selection associated with GYN/OB claims. Because this subcategory contains fewer types of claim, consequently, it can yield a more accurate categorization of risk type. In addition, GYN/OB claims are an important type of claim in the AKAM program, accounting for about 33% of all claims.

given claim frequency and pre-existing conditions, different risk types (chronic versus acute diseases) would have extra predictive power for renewals.

***Hypothesis 2*: ** In an insurance market characterized by adverse selection, the claim date for normal and prolonged delivery should be frontloaded in the policy year.

In addition, to further scrutinize GYN/OB claims for clues about adverse selection, I separated all claims associated with normal delivery and prolonged delivery and charted the duration from the date of policy commencement to the birth date of a baby.[70] A pregnancy usually lasts for forty weeks (roughly nine months). My hypothesis was that if the policy has been in effect for less than seven months at the time of birth, I could assume the policyholder was aware of her pregnancy at the time of choosing the coverage. This was defined as a "known pregnancy." If the duration was between 10 and 12 months, then most likely the policyholder was not pregnant at the time of enrollment. This was then defined as a "random pregnancy." If the duration was between eight and nine months, it was not definitive whether or not the policyholder knew of her pregnancy at the time of enrollment, and the situation could be a mixture of "known pregnancy" and "random pregnancy." The same rules applied to prolonged delivery.

I then compared the frequency of known and random pregnancies to see if most of the claims were frontloaded within the first seven months of a policy year. If this were true, then it shows adverse selection in maternity-related claims. Because AKAM does not require a waiting period for any benefit, including maternity benefits, adverse selection could be rampant. Normal delivery and prolonged delivery are major types of claims that together account for 22% of all

[70] I only included claim data from November 2008 and 2009 and the July 2009 enrollment periods. Data from the July 2010 enrollment period was excluded because it contained only the first five months of claim records.

claims.[71] I further compared the crude birth rate within this program with that in Pakistan in general to secure another piece of comparative data.

5.3 Empirical Results

5.3.1. Results for Hypothesis 1

Using data from the AKAM micro health insurance program in Pakistan, I analyzed how a household's risk type affects its decision to renew a policy.[72] Table 9 shows the results from models 1 and 2.

The results from both models are consistent using both OLS and logit regression. In Model 1, I ran the regression with all households that filed at least one claim in the November 2008 and July 2009 enrollment periods. It shows that all things being equal, a medium-risk type household was 6% more likely to renew its policy, compared with households with only acute disease claims. Both low- and high-risk households also tended to have a higher propensity to renew (5% and 2%, respectively), although these effects are not statistically significant.

Similar to the results with Model 1 in Chapter 4, households that had previously renewed were, all things being equal, 14% more likely to renew than households that had not previously renewed. Each additional claim filed also increased the probability of renewing by 5%.

[71] In this analysis I only included normal delivery and prolonged delivery claims out of GYN/OB claims, so the percentage drops from 33% to 22% of all claims.
[72] Please refer to 3.2.2 Claim data risk type categorization for a detailed explanation.

TABLE 9

Regression Results for Model 1 (all Claim Records) and Model 2 (HH with GYN/OB Claim)

Dependent variable=1 if HH is going to renew for the next period				
	Model 1		Model 2	
Variable	OLS	logit	OLS	logit
Low risk	0.05	0.05	0.05	0.05
	(0.117)	(0.142)	(0.326)	(0.372)
Medium risk	0.06**	0.06**	0.10	0.11
	(0.046)	(0.050)	(0.136)	(0.136)
High risk	0.02	0.02	0.08**	0.08**
	(0.436)	(0.449)	(0.025)	(0.026)
Previously renewed	0.14***	0.14***	0.15***	0.15***
	(0.001)	(0.001)	(0.005)	(0.004)
Frequency	0.04***	0.05***	0.04***	0.05***
	(0.001)	(0.001)	(0.001)	(0.002)
Pre-existing	0.03	0.03	-0.004	-0.003
	(0.276)	(0.275)	(0.936)	(0.952)
Log (family size)	0.10***	0.11**	0.11**	0.12**
	(0.001)	(0.016)	(0.014)	(0.016)
Head age	0.0005	0.0006	0.002*	0.002*
	(0.472)	(0.457)	(0.092)	(0.084)
Head male	0.01	0.01	-0.07*	-0.08*
	(0.561)	(0.587)	(0.067)	(0.064)
Fraction child	-0.002	0.001	0.012	0.013
	(0.971)	(0.985)	(0.875)	(0.886)
Fraction teenager	-0.09	-0.10	-0.21**	-0.23**
	(0.104)	(0.106)	(0.022)	(0.021)
Fraction elderly	0.07	0.08	-0.15	-0.16
	(0.205)	(0.219)	(0.198)	(0.208)
Fraction female	-0.12**	-0.13**	-0.05	-0.06
	(0.017)	(0.016)	(0.493)	(0.493)
July 2009	-0.15***	-0.16***	-0.13***	-0.15***
	(0.001)	(0.001)	(0.001)	(0.001)
LSO fixed effect	X	X	X	X
R square	0.61		0.54	
AIC	4,188	3,982	2,074	1,970
N	3,086	3,086	1,507	1,507

Note: The dependent variable is a dummy variable indicating whether a household renewed in the next period, and I ran both OLS and a logit model (marginal effects reported) using two specifications: Model 1 with all households with claims in the November 2008 and July 2009 periods, and Model 2 with only households that filed at least one GYN/OB claim in the November 2008 and July 2009 periods. P values are shown in parentheses.

*significant at 10%; **significant at 5%; ***significant at 1%

As for household characteristics, family size had a positive impact on the probability of renewing. Each 10% increase in family size contributed to a 1.1% increase in a household's propensity to renew a policy. Among all variables regarding age and gender composition of households, only the female ratio had a significant negative impact on willingness to renew. Each 10% increase in the female ratio decreased the probability of renewing by 1.3% when all other elements were constant.

In Model 2, when the sample was restricted to those households that filed at least one GYN/OB claim, the results of the risk type dummy variables were consistent by and large with Model 1. Again, all three variables for risk types were positively correlated with the probability of renewing (5%, 11%, and 8%, respectively, for low-, medium-, and high-risk types of households), but the result was only significant for households classified as high risk. Compared with households that filed only acute disease claims, high-risk households, all other things being the same, were 8% more likely to renew their policies.

We observed similar results with regard to previous renewals, claim frequency, and family size, all of which increased the probability of renewing.

5.3.2. Results for Hypothesis 2

I ran additional analyses for all claims associated with normal and prolonged deliveries for two reasons. First, normal delivery of babies is the most prominent type of claim in this program, and these two types of claims jointly account for a large portion of claims (22% in terms of claim frequency). Second, pregnancy is not a purely random occurrence. In a program like AKAM with no individual underwriting or waiting periods for benefits, it is very likely that pregnant women chose to enroll in the program to cover maternity costs. The cost of a normal delivery is 2,500 PKR, and the premium per person is 400 PKR. Given these numbers, in theory, if a pregnant woman lived in a household with no more than six people, enrollment was a guaranteed "good deal."

All deliveries were divided into three categories, as noted earlier, of known, random, and undetermined pregnancies. Figure 4 clearly shows a descending trend in claims for both normal and prolonged deliveries as the policy year progressed. Claim frequency was much higher in the first seven months of the policy year. On average, the monthly frequency was 211 claims for the first seven months, versus a mere 85 claims per month in the last three months of the policy year. Overall, the total delivery cases filed in the first seven months count for 75% of the total delivery cases, while the claims filed during the last three months altogether count for 13%. Figure 5 depicts the same claim duration for acute disease claims in the dataset. Unlike the results for deliveries, claims for acute disease did not show a sharp decrease in frequency as the policy year progressed. Therefore, the contrasting trends in figures 4 and 5 show the existence of adverse selection in maternity-related claims.

FIGURE 4

Distribution of Normal Delivery and Prolonged Delivery Claim Duration in a Policy Year[73]

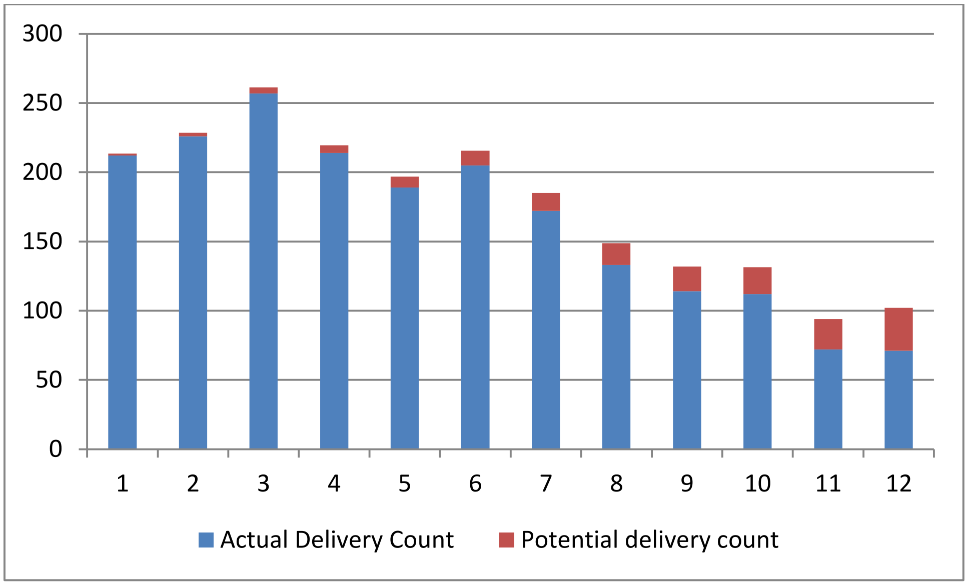
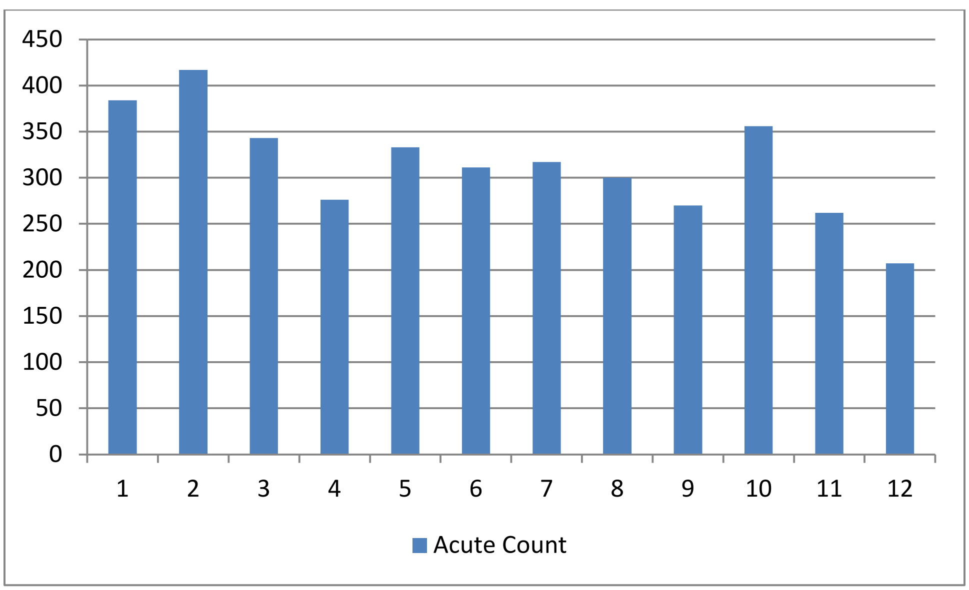


FIGURE 5

Distribution of Acute Disease Claim Duration in a Policy Year[74]



[73] For this analysis, I only included claim data from the November 2008, July 2009, and November 2009 enrollment periods. I excluded claim data from the July 2010 enrollment period because only the first five months of claim data was available for that period. Including it would have biased the duration analysis.

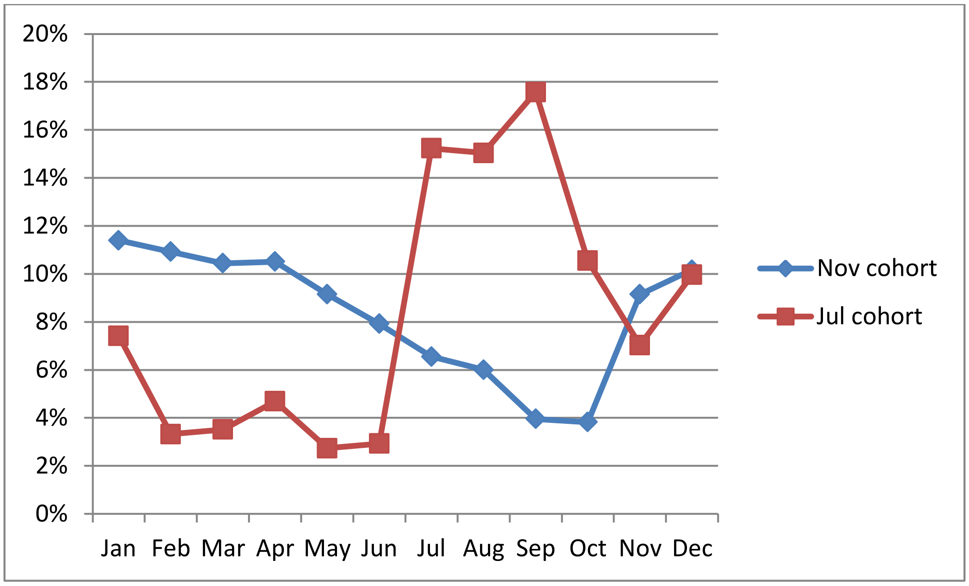
[74] The enrollment periods are the same as in the previous analysis.

However, this assertion could be countered by an argument that the descending trend of claims is a natural result of the exhaustion of insurance coverage as time goes by instead of being the result of adverse selection. To show that this was not the case, the red bars in Figure 4 graph the number of potential deliveries, which is the number of women of child-bearing age[75] who had a low balance of insurance coverage left in each month of the policy year. The average cost of a normal delivery is 2,500 PKR, and I used 3,000 PKR remaining within their insurance coverage limit as the benchmark to define "low balance." If all the women of child-bearing age with a small amount of insurance remaining were presumed to have borne a baby during the policy year and the birth dates were distributed evenly across the year, this potential delivery count is marked with a red bar. Despite this overestimated potential delivery count, the claim trend still declines as it moves toward the end of the policy year. The average monthly total delivery count, which is the sum of the actual delivery count and the potential delivery count, was 217 for the first seven months of the policy year, versus 109 for the last three months. This analysis demonstrates that when the effect of the limits of insurance coverage is considered, it is still obvious that on average, known pregnancies exceeded random pregnancies, which means that people select into the program for its delivery coverage.

[75] Defined as age 15 to 60 in this analysis.

FIGURE 6

Distribution of Normal Delivery and Prolonged Delivery Claim Duration by Enrollment Period



To further show that there is no seasonality in birth months that would contribute to the trend shown in Figure 4, I divided claim data into the November and July enrollment cohorts for the claim duration analysis. The x-axis shows the month that the claim was filed, and the y-axis records the percentage of delivery cases in that month of all the delivery cases filed within that enrollment period. Each of the July and November cohorts shows a descending trend in its own timeline, and the two cohorts display no consistency in the birth month trend. Therefore seasonality is not a major concern. For detailed information on the estimation of claim duration, please refer to Appendix 2.

To further support the finding that adverse selection exists in this program, Table 10 lists the comparisons of the crude birth rate, defined as the annual number of births per 1,000 persons in the population, in Pakistan and within three enrollment periods in the AKAM program. It clearly shows that the crude birth rate is higher among AKAM policyholders than in the general

Pakistan population. In particular, the crude birth rate in the July 2009 enrollment period was double the rate in the rest of Pakistan in 2009 (56.27 compared to 27.62), which also reflects the existence of adverse selection in maternity-related claims.

TABLE 10

Summary of Crude Birth Rate in Pakistan and AKAM Program[76]

Year	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Pakistan	30.40	29.59	31.22	30.42	29.74	27.52	28.35	27.62	25.30	24.81
AKAM Nov 08							34.44			
AKAM Jul 09								56.27		
AKAM Nov 09								33.95		

Requiring waiting period is an effective strategy commonly used to alleviate adverse selection in developed markets as well as in microinsurance markets, and was suggested by Murray.[77] Thus, faced with significant evidence of severe adverse selection in the AKAM program, especially in maternity-related claims in which a policyholder would have a sound grasp of her healthcare needs at the moment of enrollment, one proposal to improve the loss ratio is to impose a waiting period for certain benefits.

However, there was no realistic way that I could directly observe the impact of a waiting period on this group of participants, nor is there an easy way to precisely simulate people's behavior in the face of an enforced waiting period. Therefore, any analysis of the effect of a waiting period requires some assumptions at the outset. In my approach, I designed two scenarios with vastly different assumptions with the aim of setting extreme boundaries that might capture the true story between them.

[76] http://www.indexmundi.com/g/g.aspx?c=pk&v=25
[77] Murray (2011)

In the first scenario, I assumed participants have at the moment of enrollment complete knowledge about the claims they would have filed during the waiting period, and they enroll only because they want to take advantage of their complete knowledge. In other words, the waiting period policy will rule out those families who would otherwise have filed claims during this period. In contrast, in the second scenario, I assumed the participants at the moment of enrollment have no knowledge of a prospective claim that they would have filed during the waiting period. In this case, the enrolled families would be exactly the same as the enrolled group in reality, but they will not get any reimbursement for claims made during the waiting period (the premium charge, of course, will also be trimmed proportionately). In reality, people have only partial knowledge of their healthcare needs at the time they enroll in an insurance program, so an analysis of the two scenarios above will serve as boundaries, with the true impact falling somewhere in between. [78]

To enrich this analysis, I also proposed three different lengths of waiting periods and made them applicable to two benefit scenarios. I considered waiting periods of three months, six months, and nine months in which the corresponding forward contract will be trimmed to nine months, six months, and three months. I first studied the impact of applying these waiting periods to all insurance benefits (the "All benefits" scenario in Table 11). Then I focused my attention especially on the impact of enforcing a waiting period for only a delivery benefit (the "Delivery benefit only" scenario in Table 11). I isolated the latter type of benefit because normal delivery is not only the claim with the highest frequency, but also the type of claim that is easiest to predict and plan at the moment of enrollment.

[78] Please note the difficulty, as Table 11 illustrates, of determining which scenario should always serve as an upper or lower boundary. In fact, this determination hinges on whether the families ruled out in the first scenario because of the waiting period are in reality good or bad risks in the remaining contract period. Table 11 implies that these rejected families are "bad risks" in the "all benefit" case but "good risks" in the case of the "delivery benefit only," which agrees with an intuitive assessment about the effect of the presence of adverse selection.

TABLE 11

Loss Ratio Estimate under Different Waiting Period Assumptions[79][80]

	All benefits		Delivery benefit only	
Waiting period	Scenario 1	Scenario 2	Scenario 1	Scenario 2
0 month	2.01	2.01	2.01	2.01
3 months	1.69	1.77	1.95	1.94
6 months	1.46	1.55	1.88	1.88
9 months	1.36	1.41	1.84	1.84

As a baseline, the actual overall loss ratio is 2.01 when no waiting period for any benefit is required. I saw a steadily decreasing trend in the loss ratio with the imposition of longer and stricter waiting periods. For example, under Scenario 1, a three-month waiting period before any benefit could be received would trim the loss ratio to an average of 1.69, or a 16% decrease from the current level of 2.01. With six-month or nine-month waiting periods, the loss ratio would decline to 1.46 and 1.36, respectively (27% and 32% decreases). A waiting period for a delivery-related benefit, however, would also improve the loss ratio, but only a moderate amount. Using Scenario 1 as an example, with three-month, six-month and nine-month waiting periods, the loss ratio would decline to an average of 1.95, 1.88, and 1.84, respectively, corresponding to 3%, 6%, and 8% improvements from the actual level. It is also interesting to notice the element of convexity in the diminishing loss ratio as people's knowledge of their future claims gradually weakens with lengthened waiting periods. Overall, this simple estimate shows the effectiveness that imposing a waiting period can have in improving this program's loss ratio.

Using data from the AKAM micro health insurance program in Pakistan, I analyzed how a household's risk type affected its decision to renew a policy. The preliminary analysis showed

[79] Only data from November 2008, July and November 2009 is included in this estimation. July 2010 data was excluded because only partial year claim data was available.
[80] Loss ratios are properly adjusted in response to the change in contract length.

results that indicated the existence of adverse selection. In models 1 and 2, it showed that all things being equal, the medium- and high-risk types of households were more likely to renew their policies than households with only acute disease claims. Additional analysis regarding the claim duration of normal and prolonged deliveries also showed that there are more known pregnancies at enrollment than random pregnancies. This means that people selected into the program to take advantage of its coverage. Taken together, these results suggest the existence of adverse selection.

5.4 Conclusion

In summary, my research has resulted in several strands of evidence that demonstrate the existence of adverse selection in the AKAM program. First, I found that compared with households that filed only acute disease claims, those that filed claims for chronic diseases were more likely to renew their policy. Second, this trend is particularly apparent in GYN/OB claims, showing that households with low risks (such as normal delivery claims) were less likely to renew, but those considered high risk (such as those with fetal distress claims) were more likely to renew. Third, within normal and prolonged delivery claims, a clear trend of adverse selection dominating this type of claim was readily apparent. The majority of these types of claims took place within the first seven months of the policy year, showing that women entered the insurance program knowing that they were pregnant.

This evidence of adverse selection should warn microinsurers of the necessity to apply further guidelines to prevent its occurrence. For example, an appropriate waiting period in the case of maternity-related claims is recommended to offset this effect and to increase the sustainability of a micro health insurance program.

Combining the results from chapters 4 and 5 shows that despite the existence of adverse selection and the fact that households with claims were more likely to renew their policies, the renewed households have a much improved claim frequency and total claim amount after their first year of coverage. The prevalence of pre-existing conditions among newly enrolled households that had access to affordable healthcare for the first time may explain these first-year claims. In addition, some selection issues also contribute to the high claim frequency in the first year, such as cases of "known pregnancy" among newly enrolled households. Taking these factors together, the loss ratio for renewing households actually decreases dramatically after the first year of coverage.

The overall suggestion for a microinsurer is twofold. First, a microinsurer must take the necessary steps to alleviate adverse selection in order to achieve financial sustainability. Second, microinsurers should be prepared for a high loss ratio in the early years because of the unique features and circumstances in which microinsurance operates. A successful microinsurer must be farsighted about benefits instead of rushing into a short-term "fashion" business. My study shows that the renewal business turned out to have much improved loss ratios over subsequent years. A combination of dedicated efforts to retain, or even enhance, renewal ratios and methods tailored to reduce adverse selection has the potential in the long run to improve overall risk portfolios and lead to a more sustainable business.

This study suggests two interesting questions for future research. First, although the loss ratio drops considerably for first-time renewed households, this trend is less significant in the case of second-time renewals. It would be interesting to examine the loss ratios for further generations of renewed households to see whether this trend persists and in the process gain a more thorough understanding of the evolution of a risk portfolio. Second, the AKAM program imposes a 50% minimum enrollment rate at the VO level as an important mechanism to offset adverse selection. An informative question for future research would be to determine the effectiveness of such a requirement by linking the VO level enrollment rate to the actual claim experience of each VO to check whether those VOs with higher rates of enrollment file fewer claims on average. The results could show whether adverse selection exists at the VO level and assess whether the required 50% minimum VO enrollment rate is sufficient.

Appendix

Appendix 1: Claim Diagnosis Risk Type Categories

Diagnosis	Risk type	Frequency	Average cost (in PKR)
High-risk chronic disease			
Ectopic pregnancy	4	6	23,755
Fetal distress	4	321	18,173
Osteomyelitis, chronic	4	4	17,397
Goiter	4	4	16,734
Premature rupture of membranes	4	9	15,264
Obstruction, delivered	4	6	14,720
Cholecystitis	4	21	12,085
Fetal movement, decreased	4	3	11,791
Thyroiditis	4	12	11,785
Malignant neoplasms	4	15	11,499
Hydrocephalus	4	3	9,517
Pregnancy, other complications	4	197	9,326
Deep vein thrombosis	4	3	9,291
Hyperthyroidism	4	3	9,276
Liver disease, chronic	4	8	9,092
Hematuria	4	4	8,791
Renal failure, severe	4	23	8,735
Pancreatitis, light	4	4	8,113
Intestinal obstruction	4	61	7,721
Meningitis	4	10	7,599
MI, old	4	15	7,201
Pulmonary disease	4	6	6,941
Nephrotic syndrome	4	6	6,790
Medium-risk chronic disease			
MI, acute	3	67	6,694
CVA	3	55	6,645
Angina	3	31	6,430
Menorrhagia	3	15	6,425
Metrorrhagia	3	12	6,417
COPD	3	186	6,121
Arthritis	3	17	5,944
Chronic ischemic heart disease	3	71	5,902
Tuberculosis	3	6	5,818
Diabetes	3	39	5,074

Respiratory distress syndrome	3	10	4,807
Cholangitis	3	5	4,715
Sepsis	3	103	4,692
Heart failure, severe	3	5	4,579
Cardiac arrest	3	13	4,549
Heart disease	3	4	4,396
Ascites	3	3	4,213
Pelvic inflammatory disease	3	6	4,158
HTN	3	261	3,640
HTN, malignant	3	15	3,639
Hepatitis B	3	22	3,615
Jaundice	3	31	3,419
Chest pain	3	16	3,310
Low-risk chronic disease			
Osteoporosis	3	8	3,178
Respiratory arrest	2	4	3,121
Birth asphyxia	2	5	3,069
Asthma	2	214	3,033
Peptic ulcer disease	2	245	2,989
Cardiomyopathy	2	7	2,928
Arteritis, giant cell, temporal	2	3	2,912
Jaundice, newborn	2	33	2,694
Osteoarthritis	2	13	2,482
Sepsis, neonatal	2	248	2,396
Bacteremia	2	9	2,387
Hyperbilirubinemia	2	7	2,338
Anemia, chronic	2	7	2,174
Muscle weakness	2	3	2,109
Neurogenic disease	2	3	2,067
Premature labor	2	37	1,907
Gastric ulcer	2	5	1,755
Epilepsy	2	10	1,688
Back pain w/ radiation	2	7	1,616
Vertigo, central	2	15	1,615
Placenta previa, without bleeding, severe	2	6	1,613
Anemia, neoplastic disease	2	3	1,468
Dysphagia	2	4	1,340
Acute disease			
Uterus, hypertrophy	1	82	18,746
Clubbing of fingers	1	1	18,606
Hallux rigidus	1	1	17,284
Post-term pregnancy	1	3	16,432

Prolapse, uterine	1	15	16,175
Fibroid uterus	1	7	14,977
Cholelithiasis	1	92	13,701
BPH/LUTS	1	26	13,464
Cervicitis	1	9	13,420
Varicose veins	1	5	13,290
Hernia	1	108	12,520
Cystitis, acute	1	6	12,396
Appendicitis	1	124	12,190
Undescended testis	1	8	12,077
Cystocele	1	14	11,824
Hemorrhage, light	1	3	11,680
Pleural effusion	1	4	11,125
ENT	1	46	10,764
Pilonidal cyst	1	3	10,405
Anal fissure	1	9	10,196
Hydrocele	1	6	9,597
Fracture	1	64	9,583
Pulmonary contusion	1	1	9,289
Rectal disease	1	5	9,260
Hemorrhoids	1	17	9,216
Talipes equinovarus	1	2	9,048
Osteomyelitis, acute	1	2	8,812
Cyst of ovary, follicular	1	2	8,664
Ganglion	1	2	7,971
Burn	1	10	7,522
Esophagitis	1	2	6,968
Dislocation	1	9	6,121
Tonsillitis	1	88	6,060
Cellulitis	1	48	5,922
Bleeding	1	6	5,526
Labor, prolonged	1	4	5,476
Foreign body	1	8	5,267
Lymphadenitis	1	9	5,019
Pregnancy	1	10	4,648
Abscess	1	23	4,606
Measles	1	1	4,332
Calculus	1	118	4,274
Abscess of breast, postpartum	1	5	4,270
Limb anomaly	1	1	4,245
Pyelonephritis, acute	1	33	4,214
Hepatitis A	1	2	4,193
Torticollis	1	1	4,070

Benign Neoplasms Lipoma	1	2	3,980
Abortion, light	1	191	3,952
Sebaceous cyst	1	2	3,936
Tongue tie	1	2	3,917
Injury	1	87	3,868
Breast lump	1	1	3,804
Dental	1	2	3,783
Hypercalcemia	1	1	3,538
Anemia	1	63	3,390
Seizures	1	25	3,291
Sciatica	1	12	3,268
Fever	1	49	3,246
Orchitis, epididymitis	1	5	3,240
Urinary disease	1	198	3,159
Erythema nodosum	1	1	3,089
Abdominal pain and mass	1	108	3,053
Hypoglycemia, light	1	5	3,032
pneumonia	1	468	2,996
Pain	1	2	2,873
Polycythemia vera	1	1	2,764
Chickenpox	1	7	2,687
Constipation	1	1	2,687
Mastitis	1	6	2,612
Normal delivery	1	2108	2,500
Pharyngitis, acute	1	14	2,494
Dermatomycosis	1	1	2,486
Intestinal protozoa	1	2	2,476
Infection	1	1	2,432
Hyperventilation	1	1	2,396
Labyrinthitis	1	4	2,391
Croup	1	1	2,300
Spina bifida	1	1	2,278
Gastritis, light	1	999	2,232
Mumps	1	5	2,229
Dehydration	1	24	2,228
Muscle spasm	1	17	2,227
Infection, antepartum	1	558	2,167
Bronchitis, acute	1	105	2,040
SGA, newborn	1	1	2,016
Headache	1	41	1,917
Colic, infantile	1	1	1,903
Amenorrhea	1	1	1,892
Hypocalcemia, light	1	3	1,872

Ascariasis	1	4	1,810
Diarrhea	1	122	1,790
giardiasis	1	39	1,781
Vomiting, pregnancy, light	1	3	1,641
Tachycardia, paroxysmal SVT	1	1	1,622
Urticaria	1	7	1,474
Toxoplasmosis	1	1	1,452
Vertigo, benign	1	12	1,432
Scar	1	1	1,340
Aspiration, neonatal/fetal	1	1	1,284
Transient ischemic attack	1	1	1,236
Endometritis, postpartum	1	1	1,222
Post-term infant	1	1	1,150
Placenta previa, without bleeding, light	1	2	1,085
Glomerulonephritis	1	1	1,072
Vomiting	1	1	990
Poisoning	1	1	953
Post-traumatic stress disorder	1	1	828
Cord entanglement	1	1	772
Lyme disease	1	4	765
Dyspepsia	1	3	692
Hypotension	1	1	238
Unspecified			
Menopausal disorders	0	2	25,000
Periodontitis, chronic	0	1	25,000
Renal failure	0	1	25,000
PCOS	0	1	18,439
Pancreatitis, severe	0	1	13,193
Branchial cleft sinus/fistula	0	2	11,521
Testicular disease	0	2	8,214
Hepatitis C	0	2	6,005
Lumbar disc disorder w/ myelopathy	0	2	5,189
Hypoglycemia, severe	0	2	4,288
Myocarditis	0	2	3,728
Cystic fibrosis	0	1	3,603
Hepatic failure	0	2	3,577
Irritable bowel syndrome	0	1	3,500
Gastritis, severe	0	1	3,458
Atherosclerosis	0	1	3,455
Hypokalemia	0	1	3,353
Bronchitis	0	2	2,764
Polydipsia	0	1	2,565

Spinal stenosis, lumbar region	0	2	2,232
Hypocalcemia, severe	0	2	2,028
Herpes zoster	0	1	1,773
Hyperlipidemia	0	1	1,533
Hemorrhage, severe	0	1	1,466
Coagulation defects	0	2	1,220
Pre-eclampsia	0	2	1,217
Shock	0	1	987
Allergy	0	2	932
Abortion, severe	0	1	876
Vomiting, pregnancy, severe	0	1	868
Unknown diagnosis	0	8	9,059
Claim w/o diagnosis info	0	617	6,023
Total claim frequency		9683	

Appendix 2: Summary of Claim Duration

	Policy Month	1	2	3	4	5	6	7	8	9	10	11	12
Nov 08, Jul 09,	Acute Disease Count	384	417	343	276	333	311	317	300	270	356	262	207
Nov 09	Actual Delivery Count	212	226	257	214	189	205	172	133	114	112	72	71
	Low Balance Count	17	11	18	12	18	19	15	14	8	5	5	9
	Potential delivery count	1	2	4	6	8	11	13	16	18	19	22	31
Nov 08, Nov 09	Actual Delivery Count	134	149	167	160	153	154	134	116	96	88	58	56
	Low Balance Count	16	10	15	7	18	16	8	13	3	3	2	7
	Potential delivery count	1	2	4	5	7	9	10	13	14	15	16	23
Jul 09	Actual Delivery Count	78	77	90	54	36	51	38	17	18	24	14	15
	Low Balance Count	1	1	3	5	0	3	7	1	5	2	3	2
	Potential delivery count	0	0	0	1	1	1	3	3	4	5	6	8

Appendix 3: Summary of Local Supporting Organizations (LSOs) Entry Information

LSO Code	LSO Abbreviation	LSO Full Name	Region	2007 Nov	2008 Nov	2009 July	2009 Nov	2010 July
1	AK	Al-Karim Development Organization	Ghizer		Y	Y	Y	Y
2	CH	Chatorkhand	Ghizer		Y	Y	Y	Y
3	CP	n. a.	n. a.				Y	
4	DY	Danyore	Gilgit	Y	Y	Y	Y	Y
5	GN	Ganish Development Organization	Hunza/Nagar				Y	
6	GP	Gupis Rural Support Program	Ghizer		Y	Y	Y	Y
7	GR	Gojal Rural Support Organization	Hunza/Nagar		Y		Y	
8	HD	Hyderabad Rural Support Organization	Hunza/Nagar		Y	Y	Y	Y
9	KB	Karamber Local Support Organization	Ghizer		Y	Y	Y	Y
10	KD	KADO	Hunza/Nagar	Y	Y	Y	Y	Y
11	MA	Mountain Area Support Organization	Hunza/Nagar		Y		Y	
12	SD	Shundur Local Support Organization	Ghizer		Y	Y	Y	Y
13	SU	Sungum Local Support Organization	Ghizer		Y	Y	Y	Y
14	ZD	Zulfiqarabad Development Organization	Gilgit	Y	Y	Y	Y	Y

Bibliography

Akerlof, G. (1970) "The market for "lemons": Quality uncertainty and the market mechanism," *The Quarterly Journal of Economics* 84 (3): 488-500.

Amemiya, T. (1984) "Tobit models: A survey," *Journal of Econometrics* 24(1): 3-61.

Bendig, M. and Arun, T. (2011) "Enrolment in micro life and health insurance: Evidence from Sri Lanka," *The Institute for the Study of Labor (IZA)*, DP No. 5427.

Biener, C. and Eling, M. (2011) "The performance of microinsurance programs: A data development analysis," *The Journal of Risk and Insurance* 78(1): 83-115.

Biener, C. and Eling, M. (2012) "Insurability in microinsurance markets: An analysis of problems and potential solutions," *The Geneva papers on Risk and Insurance Issues and Practice* 37(1): 77-107.

Bockstal, C. (2008) HMIS in national social protection strategies: Experiences from francophone African countries, presentation in the *4th International Microinsurance Conference*.

Bolhaar, J., Lindeboom, M. and van der Klaauw, B. (2008) "A dynamic analysis of the demand for health insurance and health care," *Tinbergen institute discussion paper*, TI 2008-084/3.

Brau, J., Merrill, C. and Staking, K. (2011) "Insurance theory and challenges facing the development of microinsurance markets," *Journal of Developmental Entrepreneurship* 16(4): 411-440.

Browne, M. (1992) "Evidence of adverse selection in the individual health insurance market," *Journal of Risk and Insurance* 59(1): 13-33.

Browne, M. (2006) "Adverse selection in the long-term care insurance market," in P.-A. Chiappori and Gollier (eds.) *Competitive Failures in Insurance Markets: Theory and Evidence*, CESifo Seminar Series, Cambridge, MA: MIT Press, pp. 97-112.

CGAP working group on microinsurance (2007) "Strategies for sustainability," *Microinsurance in focus No. 4*, November 2007.

CGAP working group on microinsurance (2006) "Performance indicators for microinsurance practitioners," *Workshop report*, Oct 2006.

Chiappori, P. and Salanie, B. (2000) "Testing for asymmetric information in insurance markets," *Journal of Political Economy*, 108(1): 56-78.

Churchill, C. (ed.) (2006) “Protecting the poor: A microinsurance compendium,” the International Labour Organization and the Munich Re Foundation.

Clement, O. (2009) “Asymmetry information problem of moral hazard and adverse selection in a national health insurance: The case of Ghana national health insurance,” *Management Science and Engineering* 3(3): 101-106.

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

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

D’Arcy, S. and Doherty, N. (1990) “Adverse selection, private information, and lowballing in insurance markets,” *Journal of Business* 63(2):145.

Dionne, G., Gourieroux, C. and Vanasse, C. (2001) “Testing for evidence of adverse selection in the automobile insurance market: A comment,” *Journal of Political Economy*, 109(2): 444-453.

Dror, D., Soriano, E., Lorenzo, M., Sarol, J., Azcuna, R. and Koren R. (2005) “Field based evidence of enhanced healthcare utilization among persons insured by micro health insurance units in Philippines,” *Health Policy* 73(3): 263-271.

Dror, D., Radermacher, R. and Koren, R. (2007) “Willingness to pay for health insurance among rural and poor persons: Field evidence from seven micro health insurance units in India,” *Health Policy* 82(1):12-27.

Einav, L. and Finkelstein, A. (2011) “Selection in Insurance Markets: Theory and Empirics in Pictures,” *Journal of Economic Perspectives*, 25(1): 115-138.

Fang, H., Keane, M. and Silverman, D. (2008) “Sources of advantageous selection: Evidence from the Medigap insurance market,” *Journal of political Economy*, 116 (2): 303-350.

Finkelstein, A. and McGarry, K. (2006) “Multiple dimensions of private information: Evidence from the long-term care insurance market,” *American Economics Review*, 96(4): 938-958.

Gao, J. (2007) “Modeling individual healthcare expenditures by extending the two-part model,” working paper.

Gao, F., Powers, M. and Wang, J. (2009) "Adverse selection or advantageous selection? Risk and underwriting in China's health-insurance market," *Insurance: Mathematics and Economics,* 44(3): 505-510.

Ito, S. and Kono, H. (2010) "Why is the take-up of microinsurance so low? Evidence from a health insurance scheme in India," *The Developing Economies,* 48 (1): 74-101.

Jutting, J. (2004) "Do community-based health insurance schemes improve poor people's access to health care? Evidence from rural Senegal," *World Development* 32(2): 273-288.

Kofman, P. and Nini, G. (2004) "Asymmetric learning in insurance markets," working paper.

Kunreuther, H. and Pauly, M. (1985) "Market equilibrium with private knowledge: An insurance example," *Journal of Public Economics* 26(3): 269.

Lammers J. and Warmerdam, S. (2010) "Adverse selection in voluntary micro health insurance in Nigeria," *AIIA Research Series* 10-06.

Manning, W., Basu, A. and Mullahy, J. "Generalized modeling approaches to risk adjustment of skewed outcomes data," *Journal of Health Economics* 24(3): 465-488.

Miyazaki, H. (1977) "The Rat Race and Internal Labor Markets," *The Bell Journal of Economics,* 8(2): 394-418.

Murray, J. (2011) "Asymmetric information and countermeasures in early twentieth-century American short-term disability microinsurance," *The Journal of Risk and Insurance*, 78 (1): 117-138.

Nguyen, H. and Knowles, J. (2010) "Demand for voluntary health insurance in developing countries: The case of Vietnam's school-age children and adolescent student health insurance program," *Social Science and Medicine* 71(12): 2074-2082.

Pauly, M., Blavin, F.E. and Meghan, S. (2008) "Is there a market for voluntary health insurance in developing countries?" NBER working paper, 14095.

Pauly, M.V. (1974) "Overinsurance and public provision of insurance: The role of moral hazard and adverse selection," *Quarterly Journal of Economics*, 88(1): 44-62.

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

Rothschild, M. and Stiglitz, J. (1976) “Equilibrium in competitive insurance markets: An essay on the economics of imperfect information,” *Quarterly Journal of Economics,* 90(4): 629-49.

Securities and Exchange Commission of Pakistan, (2009) “Securities and exchange commission of Pakistan annual report 2009.”

Wang, H., Zhang, L., Yip, W. and Hsiao, W. (2006) “Adverse selection in a voluntary rural mutual health care health insurance scheme in China,” *Social Science & Medicine*, 63(5): 1236-1245.

Wilson, C. (1977) “A model of insurance markets with incomplete information,” *Journal of Economic Theory,* 16(2): 167-207.

Zhang, L. and Wang, H. (2008) “Dynamic process of adverse selection: Evidence from a subsidized community-based health insurance in rural China,” *Social Science & Medicine,* 67(7): 1173-1182.

Zheng, W. and Zhang, C. (2010) “Efficiency evaluation of China’s new rural cooperative medical system using DEA method,” working paper.