I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11751

# Adverse Selection in Low-Income Health Insurance Markets: Evidence from a RCT in Pakistan

Torben Fischer
Markus Frölich
Andreas Landmann

AUGUST 2018

IZA Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11751

# Adverse Selection in Low-Income Health Insurance Markets: Evidence from a RCT in Pakistan

**Torben Fischer**
*University of Mannheim*

**Markus Frölich**
*C4ED, University of Mannheim and IZA*

**Andreas Landmann**
*Paris School of Economics, J-PAL and C4ED*

AUGUST 2018

Any opinions expressed in this paper are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but IZA takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The IZA Institute of Labor Economics is an independent economic research institute that conducts research in labor economics and offers evidence-based policy advice on labor market issues. Supported by the Deutsche Post Foundation, IZA runs the world's largest network of economists, whose research aims to provide answers to the global labor market challenges of our time. Our key objective is to build bridges between academic research, policymakers and society.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

**IZA – Institute of Labor Economics**

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

# ABSTRACT

# Adverse Selection in Low-Income Health Insurance Markets: Evidence from a RCT in Pakistan*

We present robust evidence on the presence of adverse selection in hospitalization insurance for low-income households. A large randomized control trial from Pakistan allows us to separate adverse selection from moral hazard, to estimate how selection changes at different points of the demand curve and to test simple measures against adverse selection. The results reveal substantial selection in individual policies, leading to welfare losses and the threat of a market breakdown. Bundling insurance policies at the household or higher levels almost eliminates adverse selection, thus mitigating its welfare consequences and creating the possibility for sustainable insurance supply.

**JEL Classification:** I13, D82, O12

**Keywords:** adverse selection, health insurance, Pakistan

**Corresponding author:**
Andreas Landmann
Paris School of Economics
J-PAL
48, Boulevard Jourdan
F-75014 Paris
France
E-mail: alandmann@povertyactionlab.org

* We gratefully acknowledge financial support from the Research Center (SFB) 884 "Political Economy of Reforms" and from Project LA 3936/1-1, both funded by the German Research Foundation (DFG). We thank Luc Behagel, Sebastian Blesse, Jeffrey Clemens, Marc Gurgand, Bruce Meyer, Imran Rasul and participants of the 2015 J-Pal Europe conference, the 2016 Warwick PhD conference, the 2016 Royal Economic Society Symposium of Junior Researchers, the 1st Evidence Based Economics Summer Meeting, the EDePo Workshop, the 2016 TAPES Conference on Social Insurance, the MACIE Research Seminar at University of Marburg, the Development Seminar at the Paris School of Economics, the 12th International Microinsurance Conference, the SFB 884 Conference "Inequality and Fairness of Political Reforms," the 2017 Risk Theory Society Seminar, the 2017 ZEW Public Finance Conference, the AEL Conference 2017, the 2017 DIAL Conference, the 70th European Meeting of the Econometric Society, the CREST Microeconometrics Seminar and the 2017 EUDN Scientific Conference for helpful comments. We are grateful for excellent support in the field from Zahid Ali, Shadil Jan and Tazeemullah Khan, and for outstanding research assistance by Jan Berkes and Fawad Ejaz. We in particular thank Dr. Rashid Bajwa, Tahir Waqar, Ghulam Rasool and Mumtaz Malik Ghaffor from the National Rural Support Programme of Pakistan (NRSP) for ensuring NRSP's continued support. This experiment received IRB approval at the University of Mannheim and is registered in the American Economic Association (AEA) RCT Registry under the ID AEARCTR-0000604; https://www.socialscienceregistry.org/trials/604.

# I. Introduction

Low-income households are plagued by financial risk, and health shocks are often the most important type of unexpected events with respect to financial distress (e.g., Heltberg and Lund 2009). Insurance solutions not only promise to protect households from a poverty trap, but might also improve their long-term health and productivity. Given the deficiencies of public health systems and inefficient public health insurance in many developing countries, the potential for market-based solutions is large.[1] From an economist's perspective, however, the ability of private insurance schemes to attain efficiency depends on the extent of adverse selection. If adverse selection is present, equilibrium demand may be below the social optimum, and at worst markets might even collapse (Arrow 1963; Akerlof 1970; Rothschild and Stiglitz 1976).

The empirical debate over adverse selection in low-income health insurance is relatively recent. Some authors find evidence for more high-risk individuals selecting into health insurance (Zhang and Wang 2008; Clement 2009; Lammers and Warmerdam 2010; Yao, Schmit, and Sydnor 2015), but other studies find no evidence of adverse selection (Jütting 2004; Dror et al. 2005; Nguyen and Knowles 2010; Banerjee, Duflo, and Hornbeck 2014). Some scholars even argue that the demand for health insurance of poor households often departs from classical economic principles and is determined by community norms (Dror and Firth 2014). The evidence, however, is limited in several dimensions. First, many studies correlate uptake decisions with ex-post measures of health risk, and hence suffer from a discrimination problem between adverse selection and moral hazard (Chiappori and Salanie 2000). Those papers that use ex-ante health measures rarely show the relevance of these measures in terms of actual health events *after* insurance take-up. Second, none of these research settings allows a rigorous assessment of the welfare consequences of adverse selection. Third, there is no systematic comparison of different insurance designs regarding adverse selection and welfare.

This paper addresses these limitations by analyzing a large-scale cluster randomized control trial (RCT) on hospitalization insurance conducted in rural Pakistan. The RCT tests different insurance schemes that are randomized across more than 500 villages. We exploit baseline health measures in addition to detailed data on health events after the introduction of insurance to analyze adverse selection. Moreover, the experiment induces exogenous price variation which

[1] The Swiss Reinsurance Company estimates that the microinsurance (i.e. low-income insurance) market comprises approximately 4 billion potential customers (Swiss Re 2010). Only about 500 million people were covered under any microinsurance contract in 2013, but most of the major insurance companies currently engage in microinsurance activities to expand this market share (ILO Microinsurance Innovation Facility 2014).

enables us to estimate demand and cost curves. Identifying these curves permits us to conduct a welfare analysis similar to that of Einav and Finkelstein (2011). To the best of our knowledge, this study is the first to apply their method with experimentally controlled price variation. Finally, we test three insurance designs that are supposed to allow for different degrees of adverse selection, and conduct a comparative welfare analysis. We construct a measure for the insurers' expected reimbursement costs for each individual's inpatient expenditures based on detailed baseline health status, health history, ex-post hospitalization expenses and claim behavior.

Our results provide strong evidence that hospitalization insurance schemes for individuals suffer from adverse selection. In particular, selection becomes more pronounced with higher premium prices, creating a trade-off between cost recovery and the quality of the insurance pool. When bundling insurance policies at the household or group level, however, adverse selection is mitigated. A welfare analysis suggests that bundled policies can sustain higher quantities and lower prices than individual policies. Further, the welfare consequences of adverse selection seem less severe in relative terms for household policies.

The setup of our experiment has high relevance for the design of insurance in developing countries. Compared to insurance markets in high-income countries, contracts in the low-income context need to maintain low premiums, exhibit a simple design and keep administrative costs low. These requirements imply a limited potential for ex-ante risk screening (Brau, Merrill, and Staking 2011). In addition, providers often lack management capacity or cannot attract qualified staff, which precludes working with a portfolio of products. On the demand side, offering a single and easily understandable insurance product (pooling contract) simplifies marketing to a target group which has often been exposed to formal insurance for the first time. The drawback of policies that do not separate different risk types and that abstain from ex-ante risk screening is that they are highly vulnerable to selection. We therefore explore simple measures against adverse selection in pooling contracts which are widely applicable in low-income insurance markets (i.e. bundling individual policies on different levels). The context of our study is typical for many low-income countries. The Pakistani government spends little resources on public health care provision; there is no universal social security system; the informal sector without any access to health insurance products is large and health expenses as a consequence cause high financial stress for low-income households. These challenges are shared by many countries in Africa and Asia, underpinning the need for scalable insurance solutions.

The remainder of the paper proceeds as follows. Section II explains the approach we use to analyze adverse selection and welfare. Section III describes the context of the experiment, the insurance innovations and the hypotheses linked to their implementation. Section IV contains information about the data collection process and provides summary statistics. Section V discusses the demand for the offered insurance policies. Section VI presents empirical results on adverse selection and its welfare consequences, and Section VII concludes.

## II. Identification of Adverse Selection

The theory of adverse selection originated in the contributions of Arrow (1963), Akerlof (1970) and Rothschild and Stiglitz (1976). All these models (and many subsequent ones) hinge on the assumption that agents select into insurance policies based on their individual risk type and premium prices. In case of adverse selection, agents with the highest expected costs are those with the highest willingness to pay. This implies that the expected costs caused by the insured should always be higher than for non-insured. Further, it implies that individuals at the margin exhibit lower expected costs than the pool of already-insured individuals, creating a downward sloping marginal cost curve. Similarly, products with higher risk coverage should attract higher risk types, creating a positive correlation between coverage and riskiness of the insurance pool. From an empirical point of view, however, it is difficult to establish the presence of adverse selection due to the discrimination problem (Chiappori and Salanie 2000). An observed positive correlation between insurance coverage and loss incidences can be caused either by higher-risk individuals selecting higher coverage (adverse selection) or by higher coverage causing behavioral changes (moral hazard).

Cohen and Siegelman (2010), who summarize the empirical literature in a developed country context, describe approaches that go beyond a simple positive correlation test. These methods include exploiting dynamic claim behavior and comparing positive correlation patterns among subgroups with different potential for selection. Most of the reviewed studies on health insurance, however, only provide some form of the positive correlation test.

Another way to test for selection is to correlate *ex-ante* measures of risk, such as subjective health status or medical history before enrollment, with insurance uptake (e.g., Wang et al. 2006). Relying on ex-ante risk proxies prevents potential confounding with moral hazard, as those ex-ante risk proxies cannot be affected by the insurance status. The drawback of using ex-ante measures is the uncertainty about how they map into future costs faced by the insurer, especially in the absence of data on ex-post health events and costs. Yao, Schmit, and Sydnor (2015) discuss recent evidence from low-income health insurance markets and document several studies using ex-ante measures, but only a few link the results to actual health expenditures after the insurance choice (one exception is Banerjee, Duflo, and Hornbeck 2014). However, without reliable evidence that ex-ante proxies indeed have predictive power for ex-post costs, those studies without ex-post costs may be of little value since a lack of adverse selection found in the data could simply be an artifact of a bad proxy.

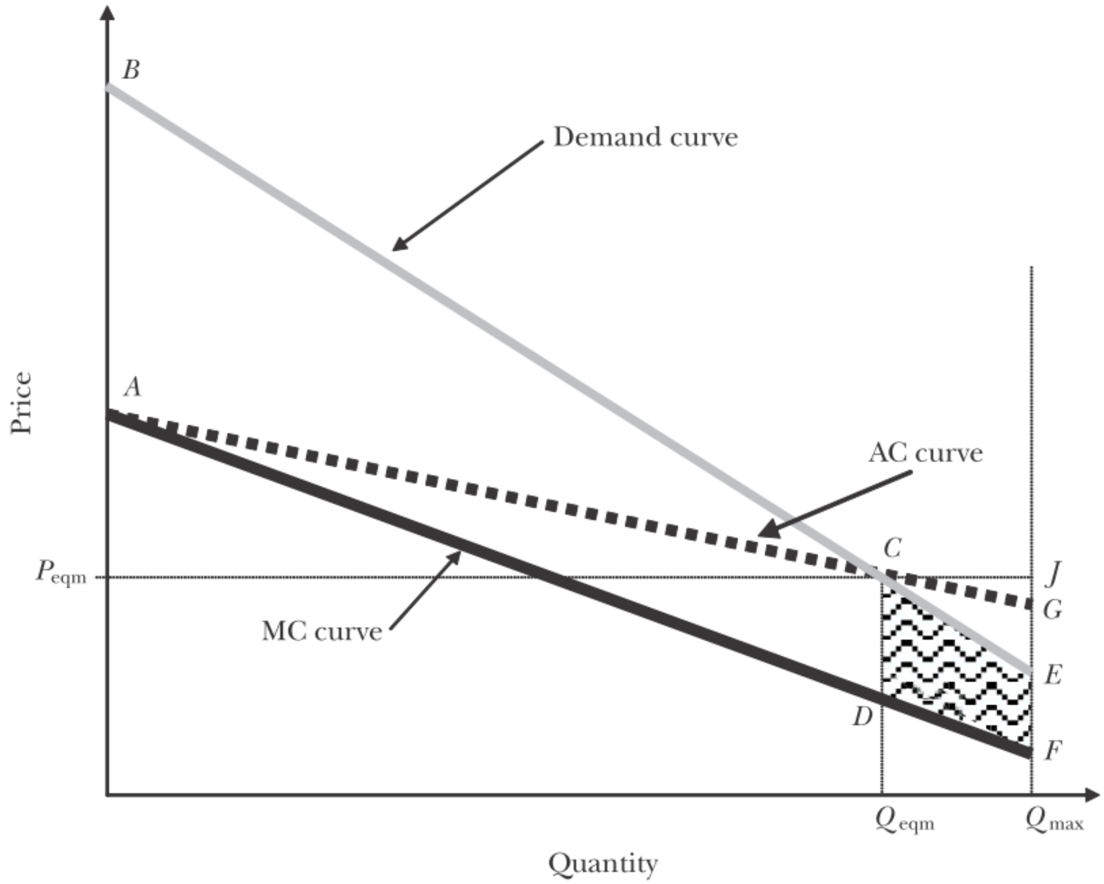
Another way to identify and quantify adverse selection is to estimate the average cost curve faced by the insurer (Einav and Finkelstein 2011). Figure 1 shows that the marginal cost curve decreases if higher-risk types exhibit a greater willingness to pay for insurance. Consequently, the insurer faces decreasing average costs with increasing demand, or adverse selection. Knowledge of the marginal and average cost curves and the demand curve not only identifies adverse selection, but also allows for welfare analyses: The intersection of demand and average cost curves determines the market allocation (under the assumption of perfect competition). A welfare loss can be observed if the willingness to pay for insurance in the market equilibrium is greater than the marginal costs of providing insurance. Figure 1 depicts this welfare loss as the shaded rectangle CDEF.

Insurance theory therefore suggests a straightforward test of the presence of adverse selection that relies on the slope of the marginal cost curve. Rejecting the null hypothesis of a flat marginal cost curve, meaning that there is no relationship between insurance price and the claim ratio, constitutes evidence for selection. Moreover, the direction of selection can be tested: A decreasing marginal cost curve suggests adverse selection, but an increasing one suggests advantageous selection.[2] The presence of moral hazard does not confound this identification approach, as the slope after the upwards shift of the average cost curve still reflects the degree of adverse selection.[3]

[2] The finding of advantageous selection would not be in line with classical adverse selection model but could result e.g., if highly risk-averse individuals purchase insurance but also take precautionary health actions, e.g., preventive health efforts, or have unobserved characteristics that also make them care about future health, which would result in the insurance-buying individuals having below-average costs.

[3] In its simplest form, moral hazard should shift the average cost curve upwards by a constant. Even in case of "selection on moral hazard" the slope still identifies adverse selection based on costs after the insurance choice, which is the most important from the insurer's perspective. This view is in line with Einav and Finkelstein (2011) who consider the selection component in moral hazard as part of adverse selection.

**Figure 1 – Analysis of Adverse Selection and Welfare**




*Source*: Figure 1 from Einav and Finkelstein (2011)

A pre-requisite for this approach is exogenous variation in the premium prices for the same insurance contract. Such exogenous variation in policy premiums allows for the estimation of demand curves while observing average costs at different demand points. Providing credible exogenous price variation, however, is usually challenging. Einav, Finkelstein, and Cullen (2010) are the first ones to use identification. They investigate the presence of adverse selection and its implied welfare costs in the context of employer provided health insurance in the US. Using countrywide data from a large US employer, they exploit differences in regional pricing to estimate both demand and average cost curves of the provided health insurance schemes. The authors find a downward sloping marginal cost curve, which constitutes evidence for the presence of adverse selection, but relatively small implied welfare cost. While several other recent studies from developed insurance markets also use the same identification approach (e.g., Hackmann, Kolstad, and Kowalski 2012; Hackmann, Kolstad, and Kowalski 2015; Finkelstein, Hendren, and Shepard 2017; Panhans 2017), we know of no studies using experimental variation in premium prices.

Within our RCT, we introduce exogenous price variation via random premium discounts. Demand and average cost curves for different insurance products can hence be estimated without any further exogeneity assumption. The costs for the insurer are calculated from ex-post health events, expenditure and claim behavior. This cost data is then used to predict expected costs for each individual based on detailed baseline health and demographic information. Predicting costs for each indvidual provides us with sufficient statistical power to compare the quality of the risk pool in different subsamples, while preserving the interpretation of the average cost curve in an expected value sense.[4] We explain the available data in Section IV, discuss the "expected cost index" in Section VI and provide more details on its construction in Appendix D.

## III. Setup of the Experiment

This section contains details on the RCT and its context. We describe the public health context in Pakistan and the role of our implementation partner in Subsection III.1. The second subsection explains the interventions as well as the most important hypotheses linked to each policy. Subsection III.3 presents our sampling strategy and the randomization procedures used for treatment allocation.

### III.1 Background

Pakistan is a lower-middle income country with a population of 189 million and a GDP per capita of USD 1,429 (2015). Almost one-third of the population lives below the national poverty line (2013).[5] Furthermore, most households are at risk of remaining or falling into poverty (World Bank 2007, 8). The government spends less than one percent of its GDP on health, which is low even for a developing country. Public health expenditure hence accounts for only about 35% of total health expenditure; 87% of private expenditure has to be paid out-of-pocket (2014). Free public health facilities exist, but service quality is perceived as low and many expensive

[4] In principle, it would be straightforward to conduct the analysis with realized claim costs only. However, hospitalization is a rare event and despite our large sample size, statistical power is too low to estimate average cost curves based on realized/reimbursed claims directly. It is especially difficult to obtain precise estimates at different demand points and for different products.

[5] See World Bank Indicators 2015 at http://data.worldbank.org/country/pakistan. Subsequent figures on public health spending and out-of-pocket expenditures are also drawn from this source.

treatments and drugs are not covered (Pakistan Ministry of Health 2009). Given the absence of a universal health insurance system, the poor are vulnerable to considerable financial risk in case of health events (Heltberg and Lund 2009). Existing schemes target public and formal sector employees, excluding the rural poor, who most often work in the informal sector. A very few NGOs and microfinance institutions offer low-income insurance policies – microinsurance – to their clients, but most of these are life insurance products bundled with a loan.

Until very recently, the National Rural Support Programme of Pakistan (NRSP), our implementation partner, was the only microinsurance provider in Pakistan offering hospitalization insurance on a significant scale (World Bank 2012, 11).[6] NRSP is the largest of 12 Rural Support Programmes in Pakistan with an outreach of more than 2.5 million households. It supports low-income households through community development activities and microfinance. NRSP is the leading provider of microcredit and the largest holder of savings among the Rural Support Programmes (Rural Support Programmes Network 2015). In rural areas, NRSP usually works with community organizations (COs), which consist of 12 to 15 member households. Members of these COs are eligible for NRSP agricultural and livestock loans that exhibit joint liability on the group level. Furthermore, NRSP offers micro-enterprise development loans to smaller, jointly liable credit groups that usually consist of three to six members.

Since 2005, NRSP has complemented its micro-credit products with mandatory hospitalization and disability insurance for its credit clients and their spouses. This policy offers three benefits. First, it covers inpatient hospitalization expenditures up to a threshold of PKR 15,000 (about USD 150) per person during the loan period. This is a significant sum relative to households' total monthly income (on average less than PKR 23,000 in our sample) and sufficient for about four days in hospital including minor surgery. Second, it separately covers accidental death and disability of the main breadwinner up to a maximum threshold of PKR 15,000.[7] Third, the outstanding loan amount is written off and a contribution of PKR 5,000 towards funeral charges is paid to the family in the case of a normal death of the main breadwinner. The annual premium of PKR 150 for both client and spouse is automatically deducted from the loan amount before disbursement. The covered expenses during hospitalization range from room charges, doctor fees, lab tests and prescribed drugs to

[6] Specific national and provincial government programs lately started to roll out similar hospitalization insurance packages in selected districts. The Prime Minister's National Health Program started in three out of 23 pilot districts until August 2016 (http://www.pmhealthprogram.gov.pk). Also in 2016, the Social Health Protection Initiative was initiated in four districts of the province Khyber Pakhtunkhwa (http://www.healthkp.gov.pk/SHPInitiative.asp).
[7] The maximal benefit depends on the degree of disability caused by the accident.

transportation costs. For maternity expenses, the reimbursement threshold is set to PKR 10,000. Pre-existing conditions are not covered. The claim process depends on the service provider. In each district, NRSP has created a panel of approved and certified hospitals. In these so-called panel hospitals, treatment expenditures up to the maximal threshold of PKR 15,000 are billed directly to the insurance company, after confirmation of the insurance status by NRSP. The patient has to cover expenses exceeding the maximal threshold. In all other facilities, the patient has to bear the medical expenses first and will be reimbursed by NRSP after approval of the claim.

### III.2 Intervention

With the insurance innovations tested in this experiment, NRSP aims to make its clients more resilient to adverse health shocks, while striving for a sustainable product. At the same time, the local context restricts the range of possible innovations. NRSP's large-scale grass-roots operations depend on simple routines and on recruiting staff from local communities. NRSP's field staff has on average nine years of formal education and its target population is mostly poor and uneducated. Any scalable insurance solution therefore needs to focus on simple contracts that are easy to administer in the field.

This study tests three simple policies that expand mandatory insurance by offering voluntary coverage for additional household dependents. A fourth policy, included in the RCT but not directly comparable to the other three designs, is also described here for completeness. The benefits and claim procedure of the offered insurance policies are similar to the mandatory insurance policy. All policies cover hospitalization expenditure and accidental death or disability up to a specific threshold. Treatment in panel hospitals is cashless up to the coverage threshold. Expenditures from non-panel facilities are reimbursed ex-post.[8]

Table 1 summarizes the insurance innovations. The *Individual* policy (P1) allows clients to enroll any number and combination of dependents. It covers hospitalization expenditures of the insured individuals up to a threshold of PKR 15,000 for a premium of PKR 100 per person insured. In addition, death or disability resulting from an accident is covered up to a maximum of PKR 15,000. The *Household* policy (P3) differs from the individual products in that the client is

[8] Further details can be taken from the terms and conditions of the insurance contracts provided in Appendix **Fehler! Verweisquelle konnte nicht gefunden werden.**. NRSP implemented a similar coverage innovation for dependents of their credit clients in Hyderabad between 2009 and 2011. This earlier pilot had promising social impacts, described in Landmann and Frölich (2015) and Frölich and Landmann (2018).

required to enroll all dependents of the household to obtain additional insurance. This policy provides the same coverage as the individual product (P1) for each insured dependent. The *Group* policy (P4) requires at least 50% uptake within the credit group or community organization. For any household of the group to be eligible, at least half of the group members present in the meeting need to enroll all their dependents. The *Individual High* policy (P2) is supposed to increase protection of clients against more expensive health events. Its coverage limits are increased to PKR 30,000 per person insured, justifying the higher premium.[9] Note that in contrast to all other schemes, the high coverage policy changes the expected reimbursement costs for a given individual and is furthermore offered at a higher price. So while the observations under this policy might help to understand how baseline characteristics translate into health behavior, the demand and claim patterns are not comparable to the other policies. We therefore focus on policies P1, P3 and P4 in our main results.

In each village, one of these four policies is offered in a community meeting. The meeting starts with an introduction to the concept of insurance and a detailed explanation of the benefits of the existing, mandatory health insurance policy. These 30- to 40-minute sessions are led by trained social organizers. Afterwards, social organizers introduce the policy which has been randomly assigned to the community. During the sign-up phase, they privately offer each client a discount voucher for 10, 20 and 30 PKR, applicable to the per-person premium for all eligible household members.

**Table 1 - Insurance Innovations**

| | Individual (P1) | Individual High (P2) | Household (P3) | Group (P4) |
|---|---|---|---|---|
| Eligibility | Individual | Individual | Household | Household |
| Add. Requirement | | | | 50% uptake in the group |
| Coverage Limit (pp) | 15,000 | 30,000 | 15,000 | 15,000 |
| Premium (pp) | 100 | 150 | 100 | 100 |
| Premium Discounts (pp) | 0-30 | 0-30 | 0-30 | 0-30 |

*Notes:* Numbers are in PKR, USD 1 ≈ 101 PKR, 15’000 PKR ≈ USD 148 (in February 2015), pp = per person.
Individual Eligibility: Client allowed to insure any number and any combination of dependents.
Household Eligibility: Client has to insure either all or none of the dependents.
Premium Discounts: Discount vouchers of 0, 10, 20 and 30 PKR (pp) were randomized with equal probability at the household level.

In terms of hypotheses, we expect a high level of adverse selection in the individual policy (P1), as clients can cherry-pick insurance coverage for high-risk household members. Compared

[9] About 80% of claims from the mandatory insurance in 2014 were above the coverage threshold of PKR 15,000. Based on these numbers and expected increases in reimbursements, the fair premium was estimated at PKR 150.

to individual insurance, the household policy (P3) is expected to impede selection of high-risk individuals, and the group policy (P4) additionally impedes selection of specific high-risk households. By construction, both bundled products should mitigate adverse selection (P4 even more than P3). The extent to which adverse selection is decreased depends on the clustering of health risks within households and groups, and on the extent to which clients possess and use information about aggregated financial risk at the level of these clusters.

The welfare implications of such risk bundling policies are theoretically ambiguous. On the one hand, we expect risk bundling to mitigate adverse selection, and thereby improve overall welfare. On the other hand, limiting the choice of clients could decrease welfare. Imagine, for example, that the marginal willingness to pay is above the uniform household price for some dependents and below this price for others. This implies an inefficient level of coverage under the household insurance (assuming that the von Neumann-Morgenstern axioms hold). The resulting demand might both be higher or lower than the individual product at the same price. Furthermore, liquidity constraints might be more of an issue in products P3 and P4, especially for large households, as premiums for all household members need to be paid. We assess demand, selection into the insurance policies and overall welfare effects in Sections V and VI.

### III.3 Sampling and Randomization

We chose the "revenue village" or "mouza," best described as a collection of settlements forming a village, as the level of randomization. This means that only one out of four interventions is available to clients living in the same village. We choose this level of randomization because it is small enough to allow for the required number of clusters, while being large enough to reach the optimal number of observations per cluster. Further, given the considerable distance between villages, this choice minimizes the potential for information spillovers, which could contaminate the treatment effect estimates. A map of the villages included in the experiment can be found in Appendix B.

The sampling procedure focuses on clients from groups whose loan application had been approved just before the introduction of the innovation in December 2014. This approach guarantees that the group composition and household structure are exogenous to the introduction of the innovations. Moreover, this procedure allows the coverage periods of the mandatory and extended insurance policies to overlap for most of the time. For sampling purposes, we first generate a unique order of credit applications from the timing in which they appear in NRSP's management information system. In a second step, we select *all* members with active loans from the pool of groups for which there is at least one credit application. New groups are added following this procedure until at least 13 client households per village are sampled to achieve an optimal cluster size.[10] Sampling from incoming credit applications implies that we do not know the set of villages with incoming credit applications ex-ante. We therefore employ a permuted block randomization procedure for dynamic treatment assignment (McEntegart 2003) and stratify the treatment assignment across a set of ex-ante village characteristics.[11] Premium discounts are randomized on the household level during the sign-up procedure. The discount is determined through a lottery in which clients have to choose one of four seemingly identical cards. These discount cards are drawn with replacement, hence giving each household the same chance for each discount. The result is captured on a sign-up sheet with unique household level identifiers.

Table 2 presents the resulting allocation of treatments. There are 502 villages with 6,461 client households, each of which is allocated either to one of the four insurance innovations or to one of two control groups. The first set of control villages constitutes a pure control group where there is no intervention in addition to the usual procedures. The sampled credit groups in the second control group, labelled "Awareness," receive a standardized session in which the contract of the existing mandatory insurance for clients and spouses is explained.[12] In our analysis, we focus on the 334 villages in which the four insurance innovations have been implemented, with policies P1, P3 and P4 being of particular interest. As expected, the number of villages across treatment arms is balanced and each treatment cluster comprises an average of 13 households.

**Table 2 - Treatment Allocation**

| | Control | Awareness | P1 | P2 | P3 | P4 | Total (Policies) | Total |
|---|---|---|---|---|---|---|---|---|
| Villages | 86 | 82 | 82 | 84 | 82 | 86 | 334 | 502 |
| Groups | 283 | 230 | 268 | 266 | 252 | 264 | 1050 | 1563 |
| HHs | 1154 | 1026 | 1022 | 1083 | 1058 | 1120 | 4283 | 6463 |
| HHs Attending | 0 | 822 | 856 | 870 | 830 | 877 | 3433 | 4255 |
| Dependents (Dep.) | 4183 | 3539 | 3560 | 3920 | 3797 | 4085 | 15362 | 23084 |
| Attending Dep. | 0 | 2798 | 2981 | 3209 | 2937 | 3156 | 12283 | 15082 |

[10] In general, this translates into sampling one complete community organization per village, sometimes amended by a smaller credit group. Alternatively, we sample four to five smaller credit groups per village.
[11] More details on the randomization procedure can be found in Appendix B.
[12] This session is also conducted in the treatment villages in which an additional insurance policy is offered.

# IV. Data

To facilitate the understanding of our analyses, the data sources and the data itself are described next.

## IV.1 Data Sources

In the analysis, we combine household and individual level data from three sources. First, we use client-level information captured in our implementation partner's management information system (MIS). Second, we collect household- and individual-level data from the sample households through computer-assisted personal interviews (CAPI). Third, we augment this information with bimonthly phone surveys for the subset of households that consented in the baseline survey.

The MIS data includes unique client, group and village identifiers that we rely on in the randomization process. In addition, our implementation partner's credit procedure involves the collection of household rosters for incoming credit clients. We use these household rosters in two ways. On the one hand, it determines insurance eligibility of the dependents at the time of the insurance offer.[13] On the other hand, we incorporate these household rosters in the survey software to facilitate the survey process. Moreover, we will have access to detailed claim data for the policies. The claim data will contain information on the type of claim (hospitalization vs. accidental death/disability), the claim amount and details on the disease diagnosed.

The household survey consists of several modules capturing socio-demographic, psychological, economic, and health indicators. The health module contains individual-level information on subjective health status, history of both in- and outpatient treatments and detailed information on coping strategies. Baseline data was collected between December 2014 and March 2015 before the implementation of the intervention. External enumerators hired by the University of Mannheim collected the data. To maximize data quality, our CAPI system included both instantaneous in-field quality assurance and regular, more sophisticated data quality checks on the enumerator level.

The phone survey captures high-frequency information on health events. In general, there is a concern that information on more regular shocks such as visits to the doctor and corresponding expenditures become inaccurate for longer recall periods. To collect complete and accurate

[13] This procedure also ensures that the household structure is exogenous to the introduction of insurance.

information on health shocks, we call respondents every two months basis and ask about the health status of their household members. The phone survey captures both inpatient and outpatient events along with the costs incurred and coping strategies. The phone survey data collection covers the complete product cycle of the insurance innovation (one year).

### IV.2 Summary Statistics

Table 3 shows some summary statistics for the 4283 households in the four insurance treatment arms. The average household size reported in the baseline survey is close to six. The average number of household members for whom the take-up information can be matched is about 5.4 and the number of eligible dependents in the household is about 3.6. The average client is about 38.5 years old and about 53% of the clients are female. Most clients have no formal education. The second panel of Table 3 (a) contains economic indicators. Average monthly income of households is about PKR 22,700 (USD 220) and on average they own about 1.4 acres of land. Further, credit obligations are about three times as large as the savings stock, which amount to about 30,000 and PKR 12,000, respectively. The third panel contains household-level health indicators. In about 12% of the sampled households, at least one member had been admitted to a medical facility for inpatient treatment in the last 12 months prior to the survey. For hospitalization, average expenditure amounts to approximately PKR 37,000 per household. On average, 18% of the sampled households have heard about insurance. Sixteen percent of the dependents in the household consulted a doctor in the last month; 2% of household members had been hospitalized in the past 12 months. Part (b) of Table 3 describes data gathered via the phone survey (93% of respondents in the baseline agree to be contacted via phone). During the 12 months covered, 15% of households report that some dependents had to be hospitalized, while two-thirds of households sought outpatient treatment for some of their dependents in the last month. On the dependent level, reported inpatient and outpatient incidences are comparable to those of the baseline survey (2% and 16% respectively).

**Table 3 - Summary Statistics**

**(a) Baseline Characteristics**

| | N | Mean | SD |
|---|---|---|---|
| *Socio-Demographics - HH* | | | |
| HH Size (Survey) | 4283 | 5.99 | 2.12 |
| HH Size (Matched) | 4283 | 5.37 | 1.91 |
| Dependents (Matched) | 4283 | 3.59 | 1.87 |
| Age of Client | 4283 | 38.62 | 10.89 |
| Client Female (D) | 4283 | 0.53 | |
| Client No Education (D) | 4283 | 0.55 | |
| *Economic - HH* | | | |
| Income (month) | 4283 | 22691 | 24695 |
| Asset Index | 4283 | 0.06 | 2.42 |
| Savings | 4283 | 12085 | 67986 |
| Credit | 4283 | 30439 | 71910 |
| *Health & Insurance - HH* | | | |
| Any Inpatient (D) | 4283 | 0.12 | |
| Total Inpatient Cost | 4283 | 4445 | 24475 |
| Knows Health Insurance (D) | 4283 | 0.18 | |
| *Health – Dependents* | | | |
| Health Step (1-5) | 15361 | 4.76 | 0.63 |
| Outpatient Experience (D) | 15361 | 0.14 | |
| Inpatient Experience (D) | 15361 | 0.02 | |
| Outpatient Cost | 15361 | 609.99 | 7920.43 |
| Inpatient Cost | 15361 | 506.36 | 7520.87 |

**(b) Phone Survey**

| | N | Mean | SD |
|---|---|---|---|
| Consent (D) | 4283 | 0.93 | |
| *Health - HH* | | | |
| Any Inpatient (D) | 4283 | 0.14 | |
| Any Outpatient (D) | 4283 | 0.65 | |
| *Health - Dependents* | | | |
| Inpatient Experience (D) | 14246 | 0.02 | |
| Outpatient Experience (D) | 14246 | 0.14 | |
| Inpatient Cost | 14246 | 371.59 | 5537.91 |
| Outpatient Cost | 14246 | 702.79 | 5415.12 |

*Notes:* The table provides means and standard deviations (SD) of the respective variables. Binary variables are indicated with (D). Monetary amounts are in Pakistani rupees (PKR), where 101 PKR ≈ USD 1.

Appendix C shows the balancing tests for these (and other) characteristics. They indicate that the randomization achieved a very good balance of covariates across treatment arms. The share of the four discount types distributed during insurance rollout is not significantly different from 25%, consistent with our uniform distribution scheme. Levels of discounts, furthermore, do not seem to differ by recipient characteristics.

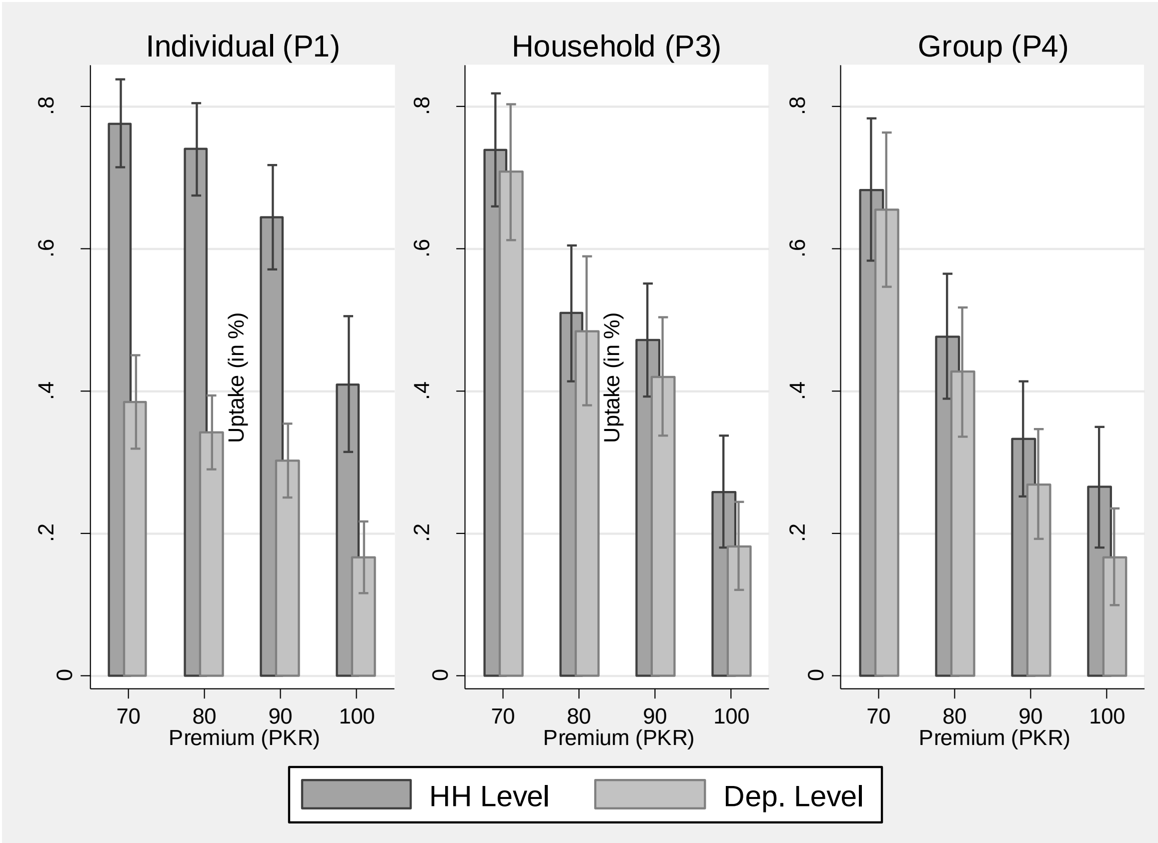
## V. Insurance Demand

Figure 2 depicts demand for the three insurance policies of interest. For each policy, demand is plotted at the four premium levels. The dark bar illustrates the share of households insuring at least one dependent, while the lighter bar illustrates the share of eligible dependents becoming enrolled in the insurance scheme.[14] All of the offered policies uptake decreases in the premium. The fraction of households covering some of their members is high in the individual policy (P1: 42-77%) compared to the household (P3: 26-74%) and the group policy (P4: 28-72%). In terms of the fraction of dependents covered, however, the bundled policies achieve higher uptake (P3: 18-71%, P4: 19-68%) than the individual policy (P1: 17-39%). Table A2 in Appendix A provides elasticity estimates assuming a linear demand curve. The resulting estimates range from -0.6 for the individual policies to -1.6 for the household policies.

In the individual product P1, we observe a large gap between the share of households and the share of individuals becoming insured at any premium level. This gap illustrates that households insure only partially. In the next section, we will analyze whether the insured individuals differ from the non-insured with respect to their expected health costs. The gap between household and individual level uptake is much lower in the household and group policies P3 and P4. This is not surprising and shows that our eligibility criteria of ensuring all dependents in the household have actually been enforced. The remaining gap exists because smaller households are more likely to purchase, again suggesting that clients struggle to insure many dependents.

[14] The figure is based on households attending the group meeting. Overall, around 80% of households attended. We do not find any statistical differences in terms of observable characteristics between households that did and did not attend the meeting (refer to Table C5 in Appendix C). The shares depicted in Figure 2 as well as the number of households attending the meetings (and their eligible dependents) are provided in Table A1 of Appendix A.

**Figure 2 - Insurance Demand, by product type**




*Notes:* The bars indicate average uptake ratios on the household and dependent level. The depicted 95% confidence intervals account for clustered standard errors at the village level. Small differences between dependent and household level uptake in policies P3 and P4 occurs because of the smaller size of insured households.

Comparing the individual policy P1 and policies with the household eligibility criterion (P3 and P4), we observe that fewer households buy insurance if enrollment of all dependents is required. However, the share of insured dependents is larger with the requirement. This suggests a trade-off between a larger pool of insured dependents and a larger pool of insured households. In other words, some households that buy (partial) insurance when offered the individual policies would not do so when they were required to insure the whole household.

Appendix Table A3 sheds further light on the determinants for households to enroll in the different insurance products. In the individual product (P1), household size does not play a role in whether to engage in some form of insurance, but larger households insure a smaller fraction of their members. Individuals selecting into the scheme tend to be in poorer health and to have a worse health history. Furthermore, children – especially the oldest son – are more likely to be enrolled. In the household and group policies (P3 and P4) individual characteristics have less predictive power. Instead, factors which might exacerbate the liquidity constraints of households (household size, female gender of the client and household experience of a hospitalization) correlate with lower take-up.

# VI. Adverse Selection

In the previous section, we estimated how many households or individuals purchase insurance as a function of the price, which exogenously varies as part of the RCT. In this section, we examine *who* purchases insurance and if these individuals systematically differ from those who do not. Thereby we analyze the relationship between insurance demand and health risk in terms of expected reimbursement costs to learn more about adverse selection.

## VI.1 Measuring Health Risk: The Expected Cost Index

Expected reimbursement costs at different demand points are of central importance for the identification of adverse selection in our setup. To measure these costs, we construct an expected cost index capturing the insurer's expected reimbursement costs for each individual given baseline covariates. To translate baseline covariates into expected costs we link characteristics to observed health events, costs and claim behavior after insurance was introduced. Even though this mapping is based on the costs observed after the introduction of the insurance innovation, the cost index remains purely a function of ex-ante characteristics.[15] We follow this approach for several reasons.

First, moral hazard can create a correlation between insurance demand and health costs after the insurance decision even in the absence of adverse selection. For example, people may change their behavior after having purchased insurance and take such behavioral changes into account before buying insurance.[16] Specifying the cost index as a function of baseline values avoids any such confounding.[17] Imagine a case where moral hazard exists and increases hospitalization costs incurred. In this case, the mapping would predict higher costs, but it would do so for all

[15] See Appendix D for further details on the parametric prediction models. Note that results are robust to other prediction models.
[16] In our case, preventive behavior may change or patients might visit more expensive facilities, both leading to an increase in the expected cost distribution of insured individuals as compared to uninsured individuals.
[17] All baseline covariates are fully exogenous in the sense that they could not be causally affected by the insurance policies offered because at the time of data collection, households were not aware of the upcoming insurance innovations. Furthermore, the household roster used to determine eligibility for insurance was collected before the innovations were introduced. Otherwise, households might have answered strategically when being asked about who belongs to their household (particularly for the household and group insurance policies P3 and P4). Table C1 reveals no statistically significant difference in the household size reported at baseline.

individuals with the same baseline variables – irrespective of their insurance status. The comparison between insured and non-insured hence remains unbiased. Note that even though the index does not suffer from a discrimination problem, our experimental setup allows further investigation of moral hazard. Specifically, we can compute predictive models for health care costs using the 162 control villages included in the RCT. Since insurance was not made available in these villages, moral hazard cannot enter into this alternative index. In contrast, estimating predictive cost models using data from the treatment villages incorporates the overall cost shift due to potential moral hazard. Appendix D reveals that both approaches lead to similar empirical results. For this reason, we regard adverse selection as the main channel, while selection on moral hazard seems less relevant in our setting.[18] The main analysis uses the predictive model that includes data from all villages in the experiment in order to maximize precision of the estimation.

Another reason to compute an expected cost index for each individual rather than using insurer reimbursement costs is that the latter relies on few claim observations. An assessment of selection across different policies and further subgroups requires a sufficient number of observations, though. Comparing individuals with respect to a large set of baseline characteristics ensures that we can effectively use all individuals for analysis and furthermore differentiate them sufficiently. A drawback of using baseline characteristics is that their interpretation is usually not trivial. Many studies employing baseline risk measures face uncertainty about how well their measures relate to the occurrence of health events in the future. Such limitations of the relevance do not apply here, as our risk measure is based on a mapping of baseline risk factors into inpatient costs arising during the product cycle. The model used for this mapping is strongly prognostic with many coefficients and the overall model being highly significant (compare Appendix D).[19]

For the main analysis below, the health risk index is computed in the same way for all individuals under the policies P1, P3 and P4, which share the same coverage limit of PKR 15,000. The average predicted cost per individual in these policies is PKR 71.42. Appendix D documents that the index is balanced between policies P1, P3 and P4.

[18] This is consistent with our expectations because the insurance covers only in-patient expenses, most of which are related to emergencies and acute illnesses, where we expect moral hazard to be less relevant.

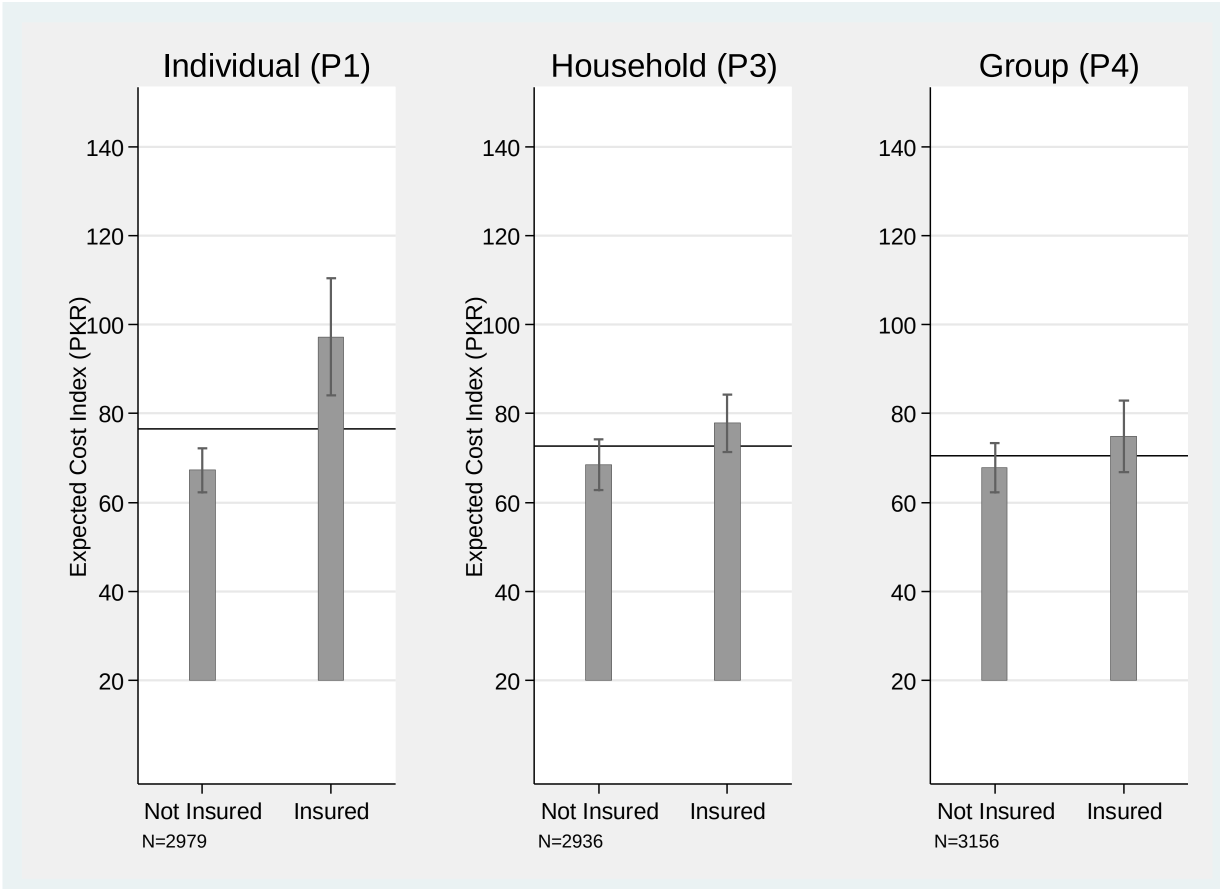
[19] Not surprisingly, the predictive power is not perfect since health shocks are unpredictable. The non-explained part reflects pure randomness as well as unobserved health risks.

### VI.2 Presence of Adverse Selection: Positive Correlation Test

As described in Section II, adverse selection leads to a situation in which high-risk types choose higher insurance coverage than lower risk types. In a first step, we therefore assess the existence of such a relationship by implementing a conventional positive correlation test (Chiappori and Salanie 2000). The individual's insurance status is given by a binary indicator for insurance uptake. Further, we proxy individuals' health risk by the expected cost index described before. Figure 3 plots coefficient estimates (and corresponding 95% confidence bounds) from a bivariate regression of the expected cost index on the binary insurance status for each of the offered policies. The horizontal line indicates the mean of the cost index. For the individual policy P1, we observe a large and statistically significant difference in the average cost index of insured versus uninsured individuals. The average cost index is almost 50% larger for insured individuals and the difference is highly significant (p-value $<<$ 0.0001). For household policies P3 and P4, we find a much smaller difference in health risk between the insured and uninsured. Average predicted costs are 10-15% higher for insured. This difference is statistically significant at the 5% level for policy P3 and insignificant for P4.

The pattern observed in Figure 3 is consistent with the presence of adverse selection. Higher-risk people are likely to become insured, especially if given a choice of individual insurance policies. The requirement to enroll all household members appears to mitigate such cherry picking and therefore might alleviate adverse selection. Note that this pattern can explain the partial insurance uptake within the household established in Section V. The corresponding demand analysis in Appendix Table A3 confirms that idiosyncratic health risk factors are a much better predictor for insurance uptake in the individual than in the household or group products. In the absence of positive assortative matching within the household, this result is mechanical in the sense that there is simply no more scope for adverse selection in the household products. At the same time, clients might be less likely to exploit the scope for selection, for example because they have difficulty obtaining an accurate estimate of the household's level of risk.

**Figure 3 - Positive correlation test: Expected cost index and take-up, by policy**




*Notes*: Bars indicate mean values of the health cost index by insurance status and policy. Confidence intervals are derived from OLS regression of the health risk index on a binary insurance status indicator. Standard errors clustered at the village level.

While this evidence of the positive correlation test seems conclusive, the behavior explaining these results remains less clear. Insurance demand is a conscious decision, but the choice might well be related to characteristics aside from expected inpatient costs. If these characteristics – such as risk aversion or income – are related to the measure of riskiness, the interpretation as deliberate selection on the basis of costs might be misleading. More risk-averse clients, for example, are expected to be more likely to insure their dependents. If these clients are at the same time more likely to be located in households with higher health risk, a result similar to that depicted in Figure 3 could arise without intentional selection based on expected costs. In appendix Table A4, we investigate this issue by explaining the demand-risk correlation with non-health related characteristics on the one hand and health history on the other. Even though the non-health variables explain some of the insurance effect, there remains a large and significant effect that can only be explained by variables related to past health events. The classical explanation for adverse selection thus appears to tell at least part of the story.

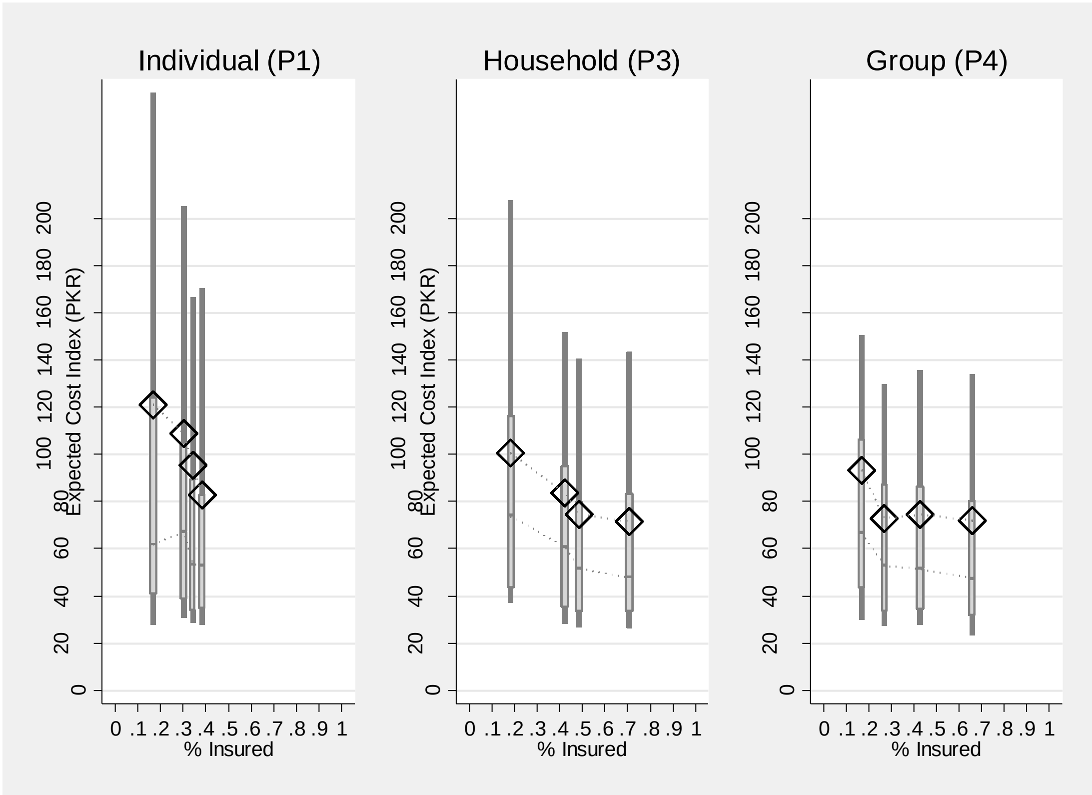
From an insurer's perspective, the behavior explaining the selection process is not the key issue. For the provider it is more interesting to know the costs of adverse selection and how these change at different levels of price and demand. Furthermore, changes in the cost distribution across prices shed additional light on the origins of adverse selection; classical explanations for adverse selection imply a decreasing average cost curve which is caused by a transition of lower-risk individuals out of the insurance pool as prices increase. The setup of our RCT allows investigating such dynamics caused by price changes. We discuss the corresponding analyses in the next section.

### VI.3 Presence of Adverse Selection: Slope of (Expected) Marginal Cost Curve

In this section we move beyond the purely correlational approach and analyze the distribution of risk types at different points of the demand curve. As illustrated in Figure 1 and discussed in Section II, the slope of the insurance providers' marginal cost curve directly indicates the presence of adverse selection (Einav and Finkelstein 2011). In the absence of adverse selection, the marginal cost curve would be flat. Thus, the risk type distribution of the insurance pool would be independent of the insurance premium. In contrast, if adverse selection were present, the marginal cost curve would be upward-sloping in price.

Figure 4 illustrates the distribution of the cost index in the pool of insured individuals at different demand levels using box plots. The box indicates the interquartile rage (IQR), with the median indicated by the line separating the box. The lower (upper) adjacent line indicates the 90$^{th}$ (10$^{th}$) percentile, respectively. The diamond represents the mean of the distribution. For the individual level policy (P1) the mean costs associated with the insurance pool decrease with demand (i.e. with lower premiums). While all depicted moments of the distribution tend to shift downward, the shift is most pronounced at the upper tail. For the household (P3) and group (P4) policy there also seems to be an upwards shift in the cost distribution with increasing premiums, but this shift is smaller than under the individual policy (P1). Table A5(a) shows the result of testing for a trend in the mean cost index of insured individuals by policy. Findings lack precision, especially when there are fewer observations in the insurance pool, but the downward slope of the average cost curve tends to be stronger in the individual policy (P1) than in the household and group policies (P3, P4).

**Figure 4 - Distribution of expected cost index of insured over demand, by policy**




*Notes*: The box plot illustrates the interquartile rage (IQR), with the median indicated by the line separating the box. The lower (upper) adjacent line shows the 90th (10th) percentile, respectively. The diamond indicates the value of the mean.

Appendix A provides further robustness checks and comparisons within the different policy regimes. Figure A1 shows the distribution of costs across demand levels amongst the non-insured. For the individual policy, there appears to be a downward shift in the cost distribution when the share of insured becomes larger. Marginal individuals switching the insurance status in response to a change in price hence seem to be high-risk relative to the non-insured but low risk relative to the insured. This is fully in line with the economic theory on adverse selection discussed in Section II. In contrast, such a pattern for non-insured is not observed under household (P3) and group (P4) policies. Table A5(b) provides a formal test for the relationship between the cost index of noninsured and the share insured. The estimated slope is significantly negative for the individual policy (P1) and insignificantly positive for household and group policies (P3, P4).

We conduct several robustness checks. For instance, we use an alternative health risk measure which is constructed by a principal component analysis of baseline health measures. Further, we repeat the analyses for the main baseline health measures separately. Our primary finding that adverse selection is much more pronounced in individual than in household and group insurance policies is robust across all these analyses.[20] Finally, we validate our analysis by comparing *real* hospitalization costs, claim incidences and claimed amounts amongst the insured during the product cycle between policy types. All three measures are significantly higher in the individual policies' insurance pool (see Table D2).

## VII. Welfare Analysis of Adverse Selection

In the previous sections we established the existence of adverse selection, especially in products for which clients can select individual members to become insured. The selection is less pronounced when complete households or groups of households have to enroll. This section investigates the welfare consequences of adverse selection under the different policies. As discussed in Section II, the exogenous price variation induced by the RCT setting identifies both the demand and the average cost curves. To analyze welfare consequences, we need to connect the demand estimates from Section V with the analyses on the slope of the average cost curve in Section VI.2. Different to the demand and cost analyses, however, we use priors to constrain our estimates to exhibit reasonable features. First, we restrict the slope of the demand curve to yield full coverage at price zero or above.[21] Second, we know that with 100% take-up, average costs of the scheme must equal the mean of the cost index in the sample. We therefore restrict the average cost curve to pass through this point. This approach is in line with the analyses in Einav, Finkelstein, and Cullen (2010). Given these restrictions, we estimate the demand curve via a linear regression of a dependent level take-up indicator on the exogenously varied premium price. The cost curve estimates result from linear regressions of the individual-specific cost index on aggregate demand for the corresponding policy at the respective price. The marginal cost curve can easily be derived afterwards in the linear case ($MC' = 2 \times AC'$). The result of the exercise is shown in Figure 5. It plots the average demand at different premium prices, the average cost index at these demand points as well as the estimated demand, average cost and marginal cost curves for the three policies. As discussed in Section II, the intersection of the demand and average cost curve determines the market equilibrium, while the intersection of demand and marginal cost curve determines the efficient allocation.

Even though the linear approximation with restrictions does not fit the data points perfectly, Figure 5 shows that sustaining insurance supply is much harder under the individual policy (P1).

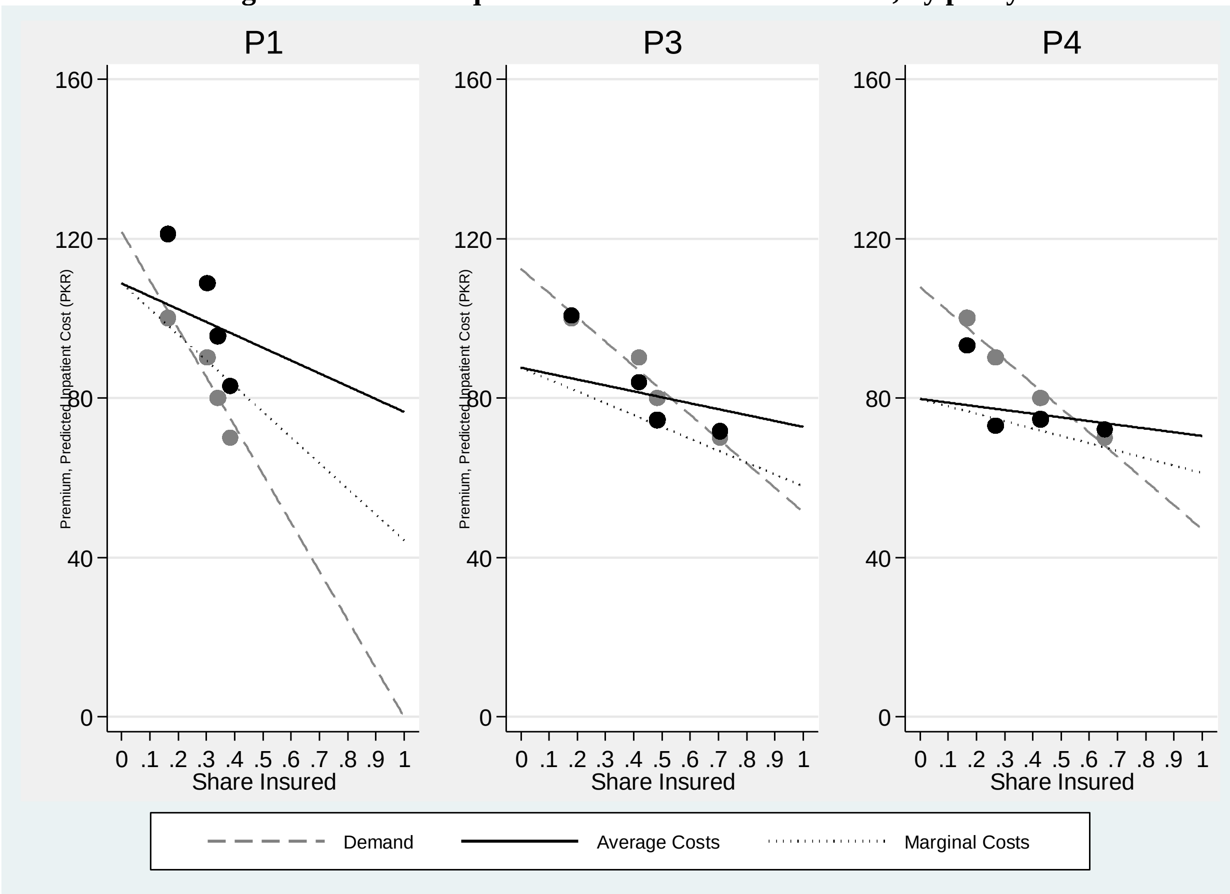
[20] The results for these robustness checks are available upon request.

[21] In other words, we assume full take-up if the product was offered for free. This restriction is binding in only one case (P1), but the fit still appears to be very good.

Both linear approximations as well as the visual inspection of data points suggest that the market for individual insurance is close to a breakdown. In the bundled policies (P3, P4), however, the average cost curves are less steep and more often situated below the demand curve. This leads to higher equilibrium demand, higher aggregate welfare and lower prices than the individual policy. This result is to some extent driven by the higher demand for insurance coverage in bundled policies (estimates shown in Table A6), but shifts in the average cost curves (parameter estimates in Table A7) also play a role. The slope of the cost curve is relatively large and highly significant for the individual policy P1 (-32.228). Figures are smaller and less significant for household policy P3 (-14.884, significant at 5% level) and group policy P4 (-9.302, insignificant). When comparing the slopes, we find significant differences between P1 and P4 (p-value: 0.0751).

It is also important to consider how close the policies are to the efficient allocation. Table 4 shows the equilibrium and the efficient allocations under the different policies and calculates the resulting welfare losses from adverse selection. Despite the lower gradient of the average cost curves for the bundled policies, losses in quantity caused by adverse selection (0.11-0.15) are higher than for the individual policy (0.09). The calculated welfare loss is also higher for the household and group insurance (P3: 1.00, P4: 0.33) than in the individual insurance policy (P1: 0.21). There are two reasons for the higher losses despite lower adverse selection in bundled policies. First, the gradient of the demand curve is lower; second, equilibrium allocations are higher. Both factors ceteris paribus extend the loss triangle. We therefore calculate the relative welfare loss, indicated in the last row of Table 4. Relative to overall welfare, losses are indeed lower in the household and group policies (10.16% and 3.50%) than in individual policy (14.40%).

**Figure 5 – Market equilibrium and efficient allocation, by policy**




*Notes*: The figure plots the demand, average and marginal cost curves for the respective policies. Average demand for the corresponding premium is given by the dots in light grey. The slope of the demand curve is estimated from a linear regression of an individual take-up indicator on the premium for which a restriction of a constant larger or equal than 1 is imposed. Average costs of the insured for the corresponding demand are given by the dots in black. The slope of the average cost curve is estimated from a linear regression of the individual level expected cost index on average take-up at the corresponding premium level. The estimation is restricted to pass through the average cost index for the respective policy at a demand level of 1. The regressions predicting the both curves are shown in Tables A6 and A7 and account for clustering of standard errors at the village level.

**Table 4 – Welfare Analysis**

| | Individual (P1) | Household (P3) | Group (P4) |
|---|---|---|---|
| *Equilibrium* | | | |
| Price | 103,41 | 79,48 | 75,02 |
| Quantity | 0,15 | 0,54 | 0,54 |
| Welfare | 1,31 | 8,84 | 8.95 |
| *Efficient* | | | |
| Price | 93,67 | 64,22 | 67,11 |
| Quantity | 0,23 | 0,79 | 0,67 |
| Welfare | 1,49 | 9,83 | 9,29 |
| *Loss* | | | |
| Quantity | 0.08 | 0.25 | 0.13 |
| Welfare | 0,18 | 0.99 | 0,34 |
| % Welfare | 11,75 | 10,06 | 3,67 |

The welfare results should be interpreted with caution, because they are sensitive to the parametric fit of the demand and cost curves. In particular, the cost estimates are based on insured individuals only and lack precision when demand is low. The restricted linear regressions smooth such fluctuations, but they also smooth away local slopes. For this reason, the quality of this parametric fit seems limited, especially for the individual policy P1. As a robustness check, we allow for a quadratic average cost curve that accounts for the analogous restriction of passing through the mean of the expected cost index at full demand. Appendix Figure A2 suggests an even more marked contrast between individual (P1) and bundled policies (P3, P4): the market for individual policies breaks down completely.[22] We therefore interpret the linear specification as a conservative estimate of the difference among policies.

Another central element of the welfare analysis is the interpretation of the demand curve. The neoclassical welfare analysis assumes that the willingness to pay measures utility derived by coverage. This interpretation could be flawed for several reasons, such as misconceptions about insurance benefits, liquidity constraints, or simply irrational behavior. We indeed find uptake patterns consistent with liquidity constraints for household and group policies (see the discussion on demand in Section V). At the same time, these findings cannot explain why demand for bundled policies is *higher* than for the corresponding individual policy. This finding of higher average willingness to pay for household than for individual insurance is not easy to reconcile with simple neoclassical theory under perfect information. In such an environment, average willingness to pay should be similar for individual and bundled policies, even though the shape of the curves might differ.[23] Finally, the interpretation of the demand curve might be distorted by the implementation of price variation through discount vouchers. Receiving a positive discount might, for example, induce more uptake than other forms of price variation. While we do not observe deviations from the linear demand predictions at particular discount levels, we cannot exclude that there are effects on demand. To severely bias our results, though, such effects would have to be different across policies.

[22] The market for individual insurance (P1) breaks down in equilibrium, even though insurance take-up would be positive in the efficient allocation. For the bundled policies (P3, P4), equilibrium prices and quantities remain similar and the equilibria are even closer to the efficient situation than in the linear specification.

[23] Assuming constant absolute risk aversion (CARA) for example, it is straightforward to show that the sum of the willingness to pay for each individual household member as indicated by the demand curve is equal to the willingness to pay for the whole household.

## VIII. Discussion and Conclusion

This paper provides robust evidence on adverse selection in low-income health insurance markets. We analyze a randomized control trial conducted in more than 500 villages of rural Pakistan in which a large local NGO offered hospitalization insurance for household members of microfinance clients. The RCT setup allows us to separate adverse selection from moral hazard, to estimate how selection changes at different points of the price curve and to test different mechanisms against adverse selection. Our analysis of adverse selection is based on individual health characteristics at baseline which we translate into an idiosyncratic expected cost index using realized costs during the product cycle.

The results suggest that there is substantial adverse selection if specific individuals within the household can be enrolled in health insurance. Adverse selection becomes worse as premiums rise, suggesting a trade-off between cost recovery and the quality of the insurance pool. Bundling policies on the household level is largely effective in mitigating adverse selection. Additional bundling of policies on the level of microfinance groups further improves the risk pool and no significant adverse selection remains in this policy.

Our main analysis assumes that the expected cost index is a good proxy to construct cost curves. An alternative and more direct approach would be to estimate average and marginal cost curves using claim data from the insurance provider only. Given that hospitalization is a rare event with a high unexplained error component, following this strategy would yield very imprecise results in our sample. Using the best predictor for expected claim costs given baseline covariates as a measure of health risk has several desirable properties in this context. It is highly relevant for expected costs, easy to interpret and at the same time its value is less affected by random health shocks at the respective price/policy points. The drawback of this measure is that we lose the selection based on health risk that is not explained by observable baseline characteristics. In that sense, results based on the cost index might represent a lower bound for true selection.

Nevertheless, the results show that (a function of) baseline health information does affect rural microfinance clients' decision about insurance uptake. Moreover, a household's ability to sort high risks into the insurance is generally limited to selection *within* households. There does not seem to be much selection on higher levels, such as the household or the micro-finance group. These findings add to the debate over classical assumptions in a developing country. While community-level demand factors might be important (Dror and Firth 2014), they apparently do not preclude microfinance clients in our sample from enrolling higher-risk members of their households.

The exogenous price variation induced in the RCT enables us to conduct a comparative welfare analysis for the different insurance schemes by merging the analyses of demand and costs curves. This exercise – which naturally rests on some assumptions – suggests that equilibrium allocations under bundled products are characterized by higher quantities, lower prices and higher welfare than under individual policies. An increased demand and decreased average cost curves under bundled policies jointly explain the result. The conclusions related to welfare are subject to some reservations, though. In addition to the difficulty to precisely identify cost and demand curves, the neoclassical assumptions needed to interpret the willingness to pay as welfare might not be fulfilled. In particular, liquidity constraints, peer effects, a lack of financial literacy or biased beliefs about future benefits could lead to uptake decisions that do not reflect the true utility derived by insurance. Furthermore, equilibrium allocations might not be relevant for a market where little supply exists so far. Irrespective of the welfare interpretation and equilibrium allocations, however, there are important observations to be drawn from the analysis. It suggests that it is easier for insurers to operate sustainably when offering bundled policies, given that the spread between willingness to pay and average costs is larger. Further, lower adverse selection under household and group policies makes entering the market less risky for insurance providers when they do not know costs and demand at specific premiums.

This paper focusses on simple pooling products. This means that only one policy is offered and no additional measures against adverse selection, such as co-payments or ex-ante screening are included. Our results show that even under these circumstances household policies might achieve a sustainable pool of insurance clients. This is good news for organizations interested in patching imperfect social security systems via insurance products for the low-income market. Such organizations might prefer a simple pooling contract to alternative solutions – such as contract portfolios with separating equilibria, screening, or risk classification based on observables – since the former are simple to market to low-income clients under difficult supply conditions and might exhibit lower administrative costs.

## References


Akerlof, George A. 1970. "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism." *The Quarterly Journal of Economics* 84 (3): 488–500.

Arrow, Kenneth J. 1963. "Uncertainty and the Welfare Economics of Health Care." *American Economic Review* 53 (5): 941–973.

Banerjee, Abhijit, Esther Duflo, and Richard Hornbeck. 2014. "Bundling Health Insurance and Microfinance in India: There Cannot Be Adverse Selection If There Is No Demand." *American Economic Review: Papers & Proceedings* 104 (5): 291--297.

Brau, James C, Craig Merrill, and Kim B Staking. 2011. "Insurance Theory and Challenges Facing The Development of Microinsurance Markets." *Journal of Developmental Entrepreneurship* 16 (4): 411.

Chiappori, Pierre-Andre, and Bernard Salanie. 2000. "Testing for Asymmetric Information in Insurance Markets." *Journal of Political Economy* 108 (1): 56–78.

Clement, Obeng Nyantakyi. 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): 101–6.

Cohen, Alma, and Peter Siegelman. 2010. "Testing for Adverse Selection in Insurance Markets." *Journal of Risk and Insurance* 77 (1): 39–84.

Dror, David M., Elmer S. Soriano, Marilyn E. Lorenzo, Jesus N. Sarol, Rosebelle S. Azcuna, and Ruth Koren. 2005. "Field Based Evidence of Enhanced Healthcare Utilization among Persons Insured by Micro Health Insurance Units in Philippines." *Health Policy* 73 (3): 263–71.

Dror, David M, and Lucy a Firth. 2014. "The Demand for (Micro) Health Insurance in the Informal Sector." *The Geneva Papers on Risk and Insurance Issues and Practice* 39: 693–711.

Einav, Liran, and Amy Finkelstein. 2011. "Selection in Insurance Markets: Theory and Empirics in Pictures." *Journal of Economic Perspectives* 25 (1): 115–38.

Einav, Liran, Amy Finkelstein, and Mark R Cullen. 2010. "Estimating Welfare in Insurance Markets Using Variation in Prices." *Quarterly Journal of Economics* 125 (3): 877–921.

Finkelstein, Amy, Nathaniel Hendren, and Mark Shepard. 2017. "Subsidizing Health Insurance for Low-Income Adults: Evidence from Massachusetts." NBER Working Paper Series. doi:10.3386/w23668.

Frölich, Markus, and Andreas Landmann. 2018. "Effects of Insurance on Child Labour: Ex-Ante and Ex-Post Behavioural Changes." *Journal of Development Studies* 54 (6): 1002–18.

Hackmann, Martin B, Jonathan T Kolstad, and Amanda E Kowalski. 2012. "Health Reform, Health Insurance, and Selection: Estimating Selection into Health Insurance Using the Massachusetts Health Reform." *American Economic Review* 102 (3): 498–501.

———. 2015. "Adverse Selection and an Individual Mandate: When Theory Meets Practice." *American Economic Review* 105 (3): 1030–66.

Heltberg, Rasmus, and Niels Lund. 2009. "Shocks, Coping, and Outcomes for Pakistan's Poor."

*The Journal of Development Studies* 45 (6): 889–910.

ILO Microinsurance Innovation Facility. 2014. "ILO's Microinsurance Innovation Facility Annual Report 2013." http://www.microinsurancecentre.org/resources/documents/unknown/ilo-s-microinsurance-innovation-facility-annual-report-2013.html.

Jütting, Johannes P. 2004. "Do Community-Based Health Insurance Schemes Improve Poor People's Access to Health Care? Evidence From Rural Senegal." *World Development* 32 (2): 273–88. doi:10.1016/j.worlddev.2003.10.001.

Lammers, Judith, and Susan Warmerdam. 2010. "Adverse Selection in Voluntary Micro Health Insuran Ce in Nigeria." 10–06. AIID Research Series. https://www.researchgate.net/publication/228428340_Adverse_selection_in_voluntary_micro_health_insurance_in_Nigeria.

Landmann, Andreas, and Markus Frölich. 2015. "Can Health-Insurance Help Prevent Child Labor? An Impact Evaluation from Pakistan." *Journal of Health Economics* 39: 51–59. doi:10.1016/j.jhealeco.2014.10.003.

McEntegart, Damian J. 2003. "The Pursuit of Balance Using Stratified and Dynamic Randomization Techniques: An Overview." *Drug Information Journal* 37 (3): 293–308.

Nguyen, Ha, and James Knowles. 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 & Medicine* 71 (12). Elsevier Ltd: 2074–82.

Pakistan Ministry of Health. 2009. "National Health Policy 2009: Stepping Towards Better Health (Draft 19 Feb 2009)." www.pc.gov.pk/Policies/Health.doc.

Panhans, Matthew T. 2017. "Adverse Selection in ACA Exchange Markets: Evidence from Colorado." https://ssrn.com/abstract=2920049.

Rothschild, Michael, and Joseph Stiglitz. 1976. "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information." *The Quarterly Journal of Economics* 90 (4): 629–49.

Rural Support Programmes Network. 2015. "Outreach #26." http://www.rspn.org/wp-content/uploads/2015/11/OUTREACH-26.pdf.

Swiss Re. 2010. "Microinsurance – Risk Protection for 4 Billion People." *Sigma* 6/2010. http://media.swissre.com/documents/sigma6_2010_en.pdf.

Wang, Hong, Licheng Zhang, Winnie Yip, and William Hsiao. 2006. "Adverse Selection in a Voluntary Rural Mutual Health Care Health Insurance Scheme in China." *Social Science and Medicine* 63 (5): 1236–45.

World Bank. 2007. "Pakistan - Social Protection in Pakistan: Managing Household Risks and Vulnerability." Washington, DC. http://documents.worldbank.org/curated/en/2007/10/8900145/pakistan-social-protection-pakistan-managing-household-risks-vulnerability.

———. 2012. "Micro Insurance in Pakistan: A Diagnostic Study." http://www.secp.gov.pk/corporatelaws/pdf/MI_Report_16102012.pdf.

Yao, Yi, Joan T. Schmit, and Justin R. Sydnor. 2015. "The Role of Pregnancy in Micro Health

Insurance: Evidence of Adverse Selection from Pakistan." *The Journal of Risk and Insurance*.

Zhang, Licheng, and Hong Wang. 2008. "Dynamic Process of Adverse Selection: Evidence from a Subsidized Community-Based Health Insurance in Rural China." *Social Science & Medicine* 67 (7): 1173–82.

**The following pages are supplementary.**

**They are intended**

**FOR ONLINE PUBLICATION ONLY.**

**They will be made available on the homepages of the authors and are included here as additional supplementary material for reviewers.**

# A. Supplementary Tables and Figures

Table A1 shows the fractions of individuals and households who bought insurance under the different insurance policies and discount levels (D0: no discount, D10: discount of 10 PKR, D20: discount of 20 PKR and D30: discount of 30 PKR). Table A2 analyzes trends and non-linearity in insurance demand.

**Table A1 - Insurance Uptake and Enforcement of Eligibility**

| | Individual (P1) | | Household (P3) | | Group (P4) | |
|---|---|---|---|---|---|---|
| | Dependents | HH | Dependents | HH | Dependents | HH |
| D0 | 0.166 | 0.415 | 0.182 | 0.258 | 0.167 | 0.265 |
| | (0.025) | (0.048) | (0.031) | (0.040) | (0.034) | (0.043) |
| D10 | 0.303 | 0.645 | 0.420 | 0.472 | 0.269 | 0.332 |
| | (0.026) | (0.038) | (0.042) | (0.040) | (0.039) | (0.041) |
| D20 | 0.341 | 0.746 | 0.484 | 0.510 | 0.427 | 0.477 |
| | (0.026) | (0.032) | (0.053) | (0.048) | (0.046) | (0.044) |
| D30 | 0.385 | 0.773 | 0.708 | 0.739 | 0.656 | 0.683 |
| | (0.033) | (0.032) | (0.048) | (0.040) | (0.055) | (0.050) |
| N | 2981 | 856 | 2937 | 830 | 3156 | 877 |

*Notes:* Standard errors in parentheses are clustered at the level of the village.

**Table A2 - Insurance Uptake and Demand Elasticities**

| | P1 | P1 | P3 | P3 | P4 | P4 |
|---|---|---|---|---|---|---|
| Premium | -0.0066*** | 0.0320* | -0.0164*** | -0.0110*** | -0.0164*** | -0.0701** |
| | (0.0013) | (0.0173) | (0.0017) | (0.0337) | (0.0020) | (0.0276) |
| Premium^2 | | -0.0002** | | -0.0000 | | 0.0003** |
| | | (0.0001) | | (0.0002) | | (0.0002) |
| Constant | 0.8636*** | -0.0.7413 | 1.8408*** | 1.6162 | 1.7726*** | 4.0090*** |
| | | | | | (0.1825) | (1.887) |
| N | 2981 | 2981 | 2937 | 2937 | 3156 | 3156 |

*Notes:* Results are from OLS regression. Standard errors in parentheses are clustered at the level of the village.

**Table A3 – Determinants of Insurance Demand by Policy**

| | Household Level Uptake | | | Individual Level Uptake | | |
|---|---|---|---|---|---|---|
| | Individual (P1) | Household (P3) | Group (P4) | Individual (P1) | Household (P3) | Group (P4) |
| *Household Level* | | | | | | |
| Discount | 0.011*** | 0.015*** | 0.015*** | 0.007*** | 0.017*** | 0.017*** |
| | (0.002) | (0.002) | (0.002) | (0.001) | (0.002) | (0.002) |
| HH Size | 0.003 | -0.048*** | -0.055*** | -0.055*** | -0.034*** | -0.038*** |
| | (0.011) | (0.010) | (0.009) | (0.006) | (0.010) | (0.008) |
| Income (in 1000 PKR) | -0.000 | -0.000 | 0.001 | 0.000 | -0.000 | 0.001 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Saving (in 1000 PKR) | 0.000 | 0.000 | 0.000 | -0.000* | 0.001 | 0.000 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Asset Index | -0.005 | 0.018* | 0.003 | 0.004 | 0.014 | 0.003 |
| | (0.008) | (0.009) | (0.008) | (0.005) | (0.010) | (0.008) |
| Head Female | 0.012 | -0.120** | -0.101* | -0.021 | -0.119** | -0.095* |
| | (0.042) | (0.047) | (0.060) | (0.033) | (0.049) | (0.055) |
| No Education | -0.060 | -0.047 | -0.061 | -0.043* | 0.004 | -0.049 |
| | (0.051) | (0.044) | (0.043) | (0.024) | (0.034) | (0.033) |
| High Education | -0.042 | -0.095* | -0.029 | -0.052* | -0.022 | 0.004 |
| | (0.050) | (0.052) | (0.057) | (0.029) | (0.028) | (0.039) |
| Any Inpatient | 0.085* | -0.022 | -0.093* | -0.012 | -0.042 | -0.082 |
| | (0.045) | (0.054) | (0.051) | (0.030) | (0.058) | (0.054) |
| *Dependent Level* | | | | | | |
| Female | | | | -0.109*** | -0.025 | -0.004 |
| | | | | (0.019) | (0.017) | (0.018) |
| Age (0-4) | | | | 0.125*** | 0.071 | 0.084 |
| | | | | (0.036) | (0.049) | (0.052) |
| Age (5-9) | | | | 0.067* | 0.049 | 0.056 |
| | | | | (0.038) | (0.045) | (0.045) |
| Age (10-14) | | | | 0.057 | -0.006 | 0.070 |
| | | | | (0.036) | (0.040) | (0.043) |
| Age (15-19) | | | | 0.061** | -0.005 | -0.003 |
| | | | | (0.030) | (0.033) | (0.031) |
| Age (20-29) | | | | ref. | ref. | ref. |
| Age (30-49) | | | | 0.038 | -0.045 | 0.037 |
| | | | | (0.042) | (0.047) | (0.039) |
| Age (50-59) | | | | 0.061 | 0.115* | 0.100* |
| | | | | (0.070) | (0.068) | (0.054) |
| Age (60-69) | | | | 0.044 | -0.019 | 0.066 |
| | | | | (0.057) | (0.060) | (0.064) |
| Age (70+) | | | | 0.112 | 0.035 | 0.168* |
| | | | | (0.082) | (0.074) | (0.092) |
| Low Health | | | | 0.183** | 0.009 | 0.013 |
| | | | | (0.083) | (0.099) | (0.089) |
| Medium Health | | | | 0.084** | -0.003 | -0.006 |
| | | | | (0.040) | (0.038) | (0.043) |
| Inpatient Treatment | | | | 0.153*** | -0.038 | -0.078 |
| | | | | (0.056) | (0.090) | (0.052) |
| Outpatient Treatment | | | | 0.066** | 0.051 | 0.003 |
| | | | | (0.032) | (0.034) | (0.034) |
| First Son | | | | 0.058** | 0.017 | 0.015 |
| | | | | (0.027) | (0.020) | (0.020) |
| First Daughter | | | | 0.027 | -0.023 | 0.034 |
| | | | | (0.029) | (0.021) | (0.023) |
| Working | | | | -0.066** | -0.029 | 0.002 |
| | | | | (0.032) | (0.028) | (0.027) |
| Constant | 0.473*** | 0.530*** | 0.476*** | 0.421*** | 0.404*** | 0.313*** |
| | (0.058) | (0.060) | (0.062) | (0.050) | (0.070) | (0.072) |
| N | 856 | 830 | 877 | 2981 | 2937 | 3156 |
| $R^2$ | 0.07 | 0.17 | 0.16 | 0.13 | 0.19 | 0.19 |

*Notes:* Point estimates result from OLS regression with standard errors clustered at the village level.

**Table A4 – Correlation between Insurance Demand and Expected Costs Index**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Controls | none | non-health covariates^ | observable by insurer~ | all |
| P1 (N=2981) | 29.927*** | 19.841*** | 19.431*** | 2.318 |
| | (6.722) | (5.582) | (5.782) | (3.121) |
| P3 (N=2937) | 9.307** | 0.291 | 1.057 | -0.323 |
| | (3.854) | (3.102) | (3.181) | (1.470) |
| P4 (N=3156) | 7.264 | -2.805 | -3.197 | 0.107 |
| | (4.793) | (3.364) | (3.347) | (1.624) |

*Notes: R*esult from OLS regression of the expected costs index on individual insurance uptake with standard errors clustered at the village level. Covariates are HH size, client gender, client education level dummy, age category dummies, HH income, HH savings, HH asset index, individual work status, individual health status, inpatient and outpatient treatment experience and related costs.

^ All variables except individual health status, inpatient and outpatient treatment experience and related costs.

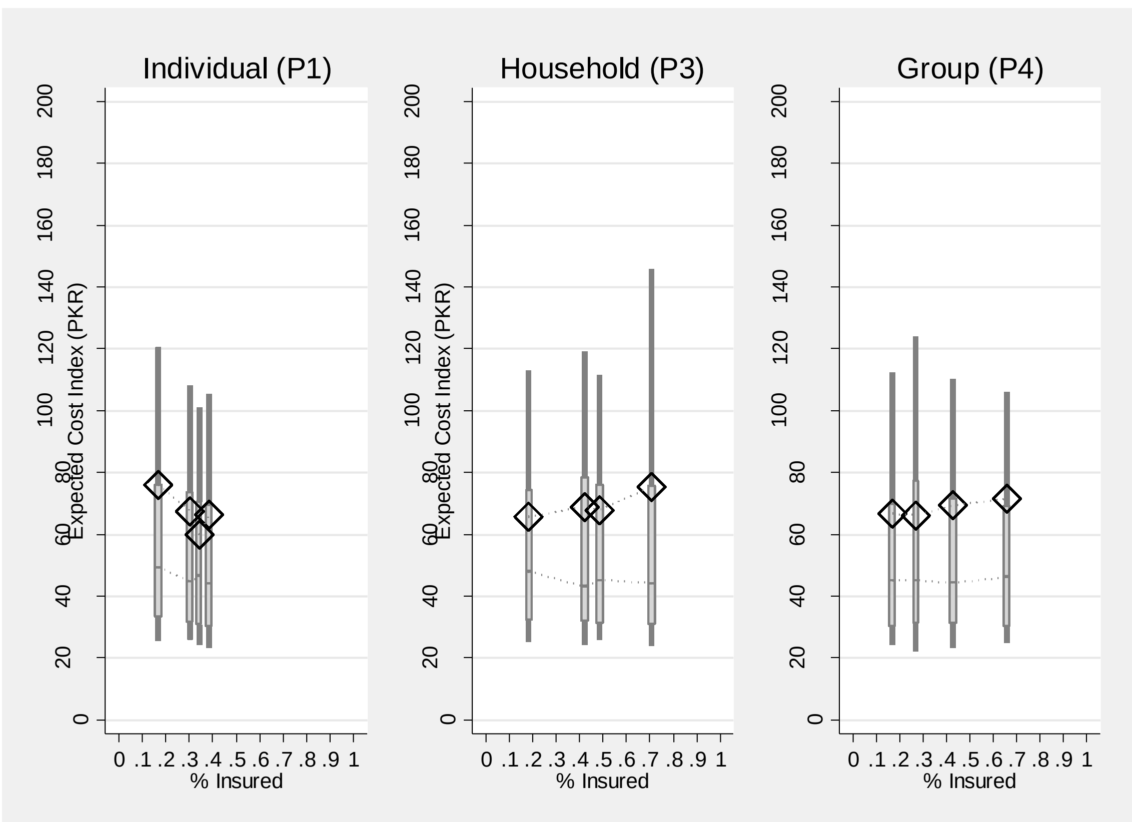
~ HH size, client gender, client education level dummy, age category dummies.

Table A4 shows the result of regressing the expected costs index on individual insurance uptake under the different insurance policies. The first specification implements a simple positive correlation test. It reveals that the difference between insured and non-insured individuals is substantially larger in the individual (P1) than in the household (P3) and group (P4) insurance schemes. Specification (2) tests whether the positive correlation can be explained by selection based on non-health factors. The idea is that the purchase decision might be influenced by non-health factors which also correlate with health risk, thus creating a positive correlation without the intention of adverse selection. Controlling for such confounding factors would therefore lead to a change in the estimated coefficient compared to the first specification. The results from specification (2) show that some of the differences between insured and non-insured individuals can indeed be explained by non-health factors. Nonetheless, most of the correlation remains in policy P1, for which the coefficient is still highly significant.

As a next step, we control for characteristics that are easy to observe and verify. The idea of this exercise is to test whether an insurance company could in principle separate risk types when using information that is available and reliable in a low-income setting under realistic conditions. Specification (3) controls for such (mainly demographic) variables. Similar to the specification before, the coefficient remains positive and significant for the individual policy (P1), suggesting that classifying individuals based on observable baseline characteristics might not solve the adverse selection problem. For illustrative purposes, specification (4) uses all control variables – essentially the ones used to create the index. As expected, the correlation disappears.

Figure A1 shows the distribution of costs across demand levels amongst the non-insured. For the individual policy, there appears to be a downward shift in the cost distribution when the share of insured becomes larger. Marginal individuals switching the insurance status in response to a change in price hence seem to be high risk relative to the non-insured but low risk relative to the insured. This is fully aligned with the economic theory on adverse selection discussed in Section II. In contrast, such a pattern for non-insured is not observed under household (P3) and group (P4) policies.

**Figure A1 - Change in risk distribution across discounts, non-insured**




*Notes:* This figure illustrates shifts in the expected cost distribution by discount level and policy regime. The box depicts the interquartile range (IQR). The middle line indicates the median. The upper (lower) adjacent line depicts the 90% (10%) quantile, respectively. The diamond indicates the mean.

Table A5(a) shows the result of testing for a trend in the mean cost index of insured individuals by policy. Findings lack precision, especially when there are fewer observations in the insurance pool, but the downward slope of the average cost curve tends to be stronger in the individual policy (P1) than in the household and group policies (P3, P4). Table A5(b) tests the relationship between the cost index and the share insured for the noninsured. The estimated slope is negative for individual policies (significant for P1), consistent with adverse selection theory, and insignificantly positive for household and group policies (P3, P4).

**Table A5 – Trend in Expected Costs**

**(a) Insured**

| | P1 | P3 | P4 |
|---|---|---|---|
| Uptake (%) | -186.817* | -47.110*** | -20.854 |
| | (100.028) | (16.653) | (13.326) |
| Constant | 158.372*** | 102.984*** | 84.670*** |
| | (34.875) | (10.534) | (7.376) |
| N | 922 | 1350 | 1211 |

**(b) Non-Insured**

| | P1 | P3 | P4 |
|---|---|---|---|
| Uptake (%) | -56.068 | 15.623 | 10.941 |
| | (34.851) | (13.745) | (17.763) |
| Constant | 84.027*** | 62.314*** | 64.043*** |
| | (10.762) | (5.567) | (5.761) |
| N | 2059 | 1587 | 1945 |

*Notes:* Point estimates result from OLS regression of expected cost index on average demand for relevant policy at respective discount, standard errors clustered at the village level.

**Table A6 – Slope of the Demand Curve, restricted**

| | Individual (P1) | Household (P3) | Household (P4) |
|---|---|---|---|
| Premium | -0.008*** | -0.016*** | -0.016*** |
| | (0.000) | (0.002) | (0.002) |
| Constant | 1.000 | 1.841*** | 1.773*** |
| | (.) | (0.161) | (0.182) |
| N | 2981 | 2937 | 3156 |

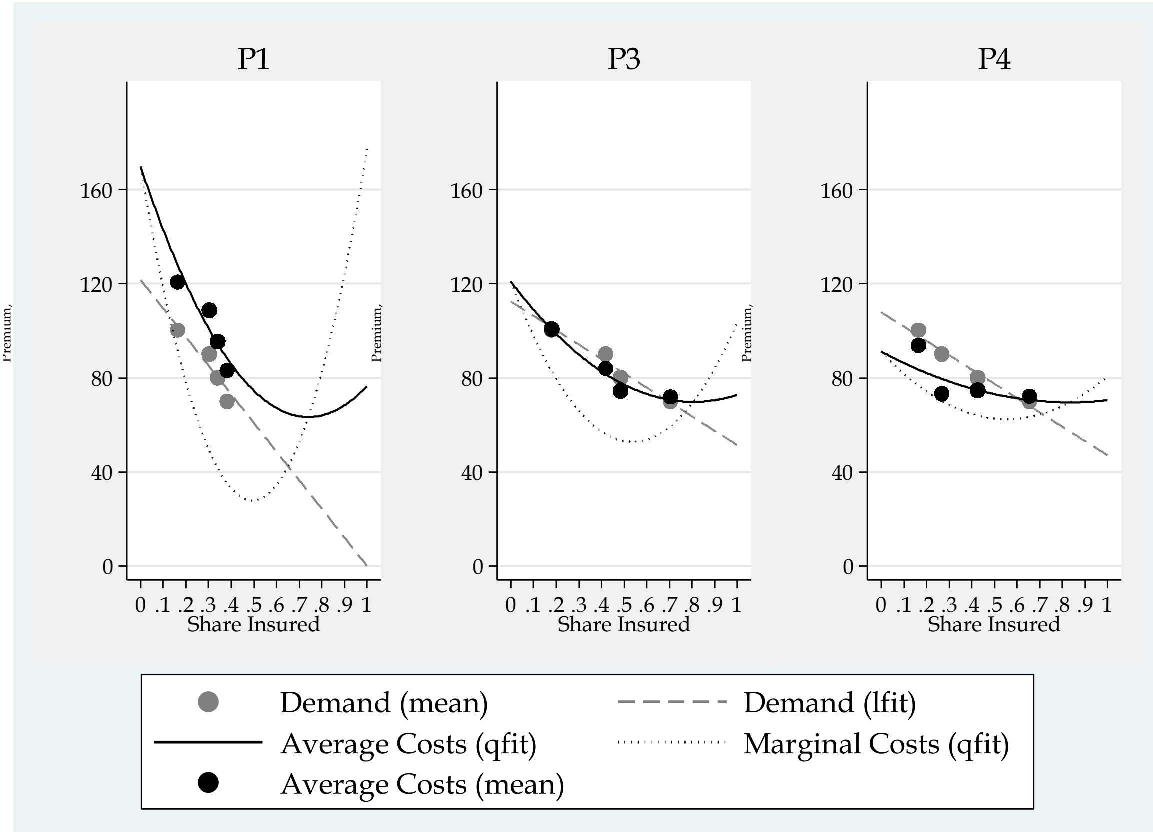
*Notes:* The slope of the demand curve is estimated from a linear regression of an individual take-up indicator on the premium, and a restriction of a constant larger or equal than 1 is imposed. Standard errors are not reported if the restriction is binding (only the case for P1). Standard errors are clustered at the village level.

**Table A7 – Slope of the Average Cost Curve, restricted**

| | Individual (P1) | Household (P3) | Household (P4) |
|---|---|---|---|
| Demand | -32.146*** | -14.841** | -9.617 |
| | (9.924) | (7.088) | (6.955) |
| Constant | 108.560*** | 87.621*** | 80.081*** |
| | (9.924) | (7.088) | (6.955) |
| N | 922 | 1350 | 1211 |

*Notes:* The slope of the average cost curve is estimated from a linear regression of the individual level expected cost index on average take-up at the corresponding premium level. The estimation is restricted to pass through the average cost index for the respective policy at a demand level of 1. Standard errors are clustered at the village level.

**Figure A2 - Market equilibrium and efficient allocation (quadratic cost curve), by policy**




*Notes*: The figure plots the demand, average and marginal cost curves for the respective policies. Average demand for the corresponding premium is given by the dots in light grey. The slope of the demand curve is estimated from a linear regression of an individual take-up indicator on the premium for which a restriction of a constant larger or equal than 1 is imposed. Average costs of the insured for the corresponding demand are given by the dots in black. The slope of the average cost curve is estimated from a quadratic regression of the individual level expected cost index on average take-up at the corresponding premium level. The estimation is restricted to pass through the average cost index for the policy at a demand level of 1.

## B. Randomization Procedure

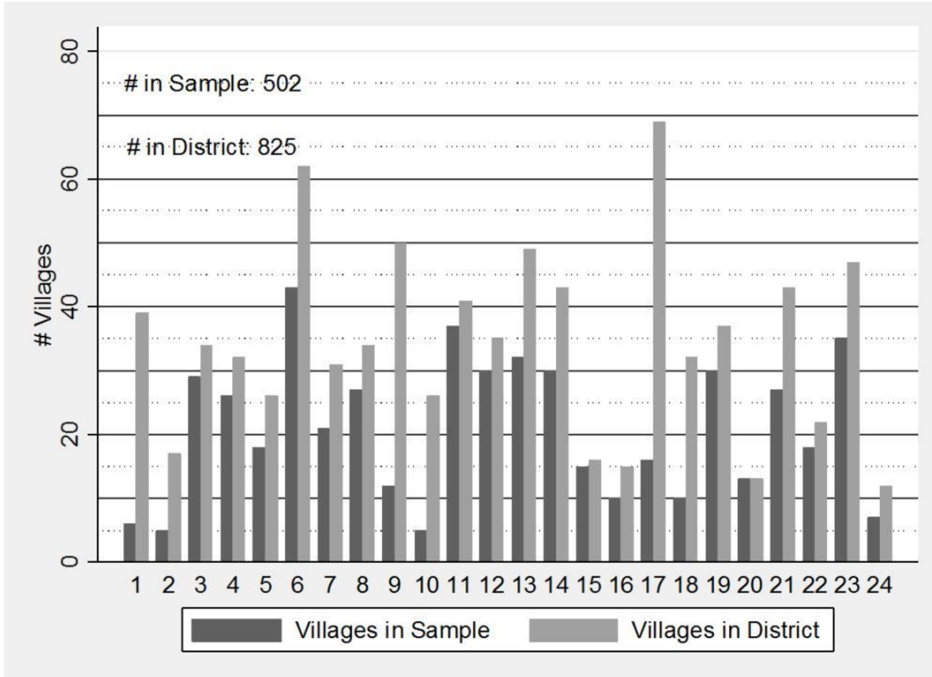
Sampling from incoming credit applications implies that we do not know the set of villages with incoming credit applications ex-ante. Instead, we start with a census of all villages in which our implementation partner operates. To achieve a balanced treatment allocation, we use a permuted block randomization procedure for dynamic treatment assignment. This procedure is used frequently in medical studies facing similar problems of patients stochastically entering the trial (McEntegart 2003). In addition, we stratify the treatment assignment across a set of ex-ante village characteristics to improve balance among treatments along a set of important characteristics.

We condition the randomization on the rural/urban status (4 categories), the historical origin of the village (2 categories) as well as the distance to the next hospital under NRSP's panel (3 categories). This leaves us with a categorization of villages into 24 strata. The treatment assignment proceeds as follows. In a first step, we generate a set of randomly permuted blocks of the six main treatment indicators for each of the 24 strata. In a second step, we produce a unique order in which the villages have entered the experiment. For this purpose, we rely on the timing of loan applications entered in the management information system (MIS). Using the list from step 2, we create strata specific lists of villages that are ordered according to the date and time they entered the MIS. In a final step, each village on this strata-specific list is matched with the corresponding treatment from the strata-specific permuted block of treatments.

This procedure guarantees a balanced distribution of treatments in each cluster, especially when there are sufficient villages per strata entering the experiment to cover full blocks. The reason is that within a full block, one village is assigned to each treatment and no imbalance can occur. Hence, the more full blocks are covered, the fewer imbalances can remain. Figure B1 shows the total number of villages in the district where the RCT takes place by strata and by the number of villages finally entering the experiment. Only three of 24 strata have fewer than six villages to create at least one full block.

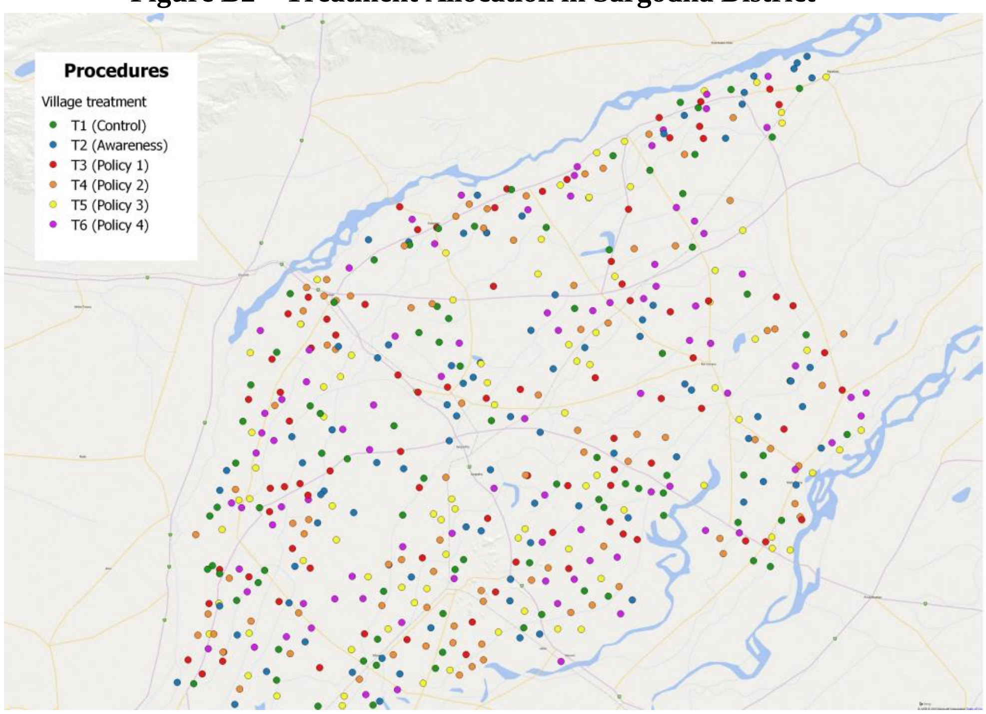
Figure B2 shows the geographical distribution of treatments. The treatment arms appear to be balanced across the whole district, suggesting that the randomization procedure worked as expected.

**Figure B1 – Distribution of clusters across strata**




*Notes:* The figures illustrates the distribution of treatment clusters across strata. The 24 strata are generated from ex-ante village level information on location (distance to closest panel hospital, 3 categories), historical origin (chak vs. no chak, 2 categories) and rural/urban status (percentage of agricultural loans, 4 categories).

**Figure B2 – Treatment Allocation in Sargodha District**




*Notes:* The figure illustrates the distribution of treatments across Sargodha district. The dots capture the center of the village. The legend gives the corresponding treatment. The average minimum distance between the villages is about 4 km and the average distance about 50 km.

## C. Balancing Tests

We present balancing tests that assess whether our randomization indeed results in a similar distribution of covariates across treatment arms. The balancing tables have the following structure. The first column shows the overall means (standard deviations are in brackets). Subsequent columns provide means and standard deviations for each treatment arm separately. The final column contains the p-value from a joint test for model significance from the following estimation equation:

$$X_{iv} = \alpha + \beta_2 I_{\{T_{iv}=P2\}} + \beta_3 I_{\{T_{iv}=P3\}} + \beta_4 I_{\{T_{iv}=P4\}} + d_s S_v + \varepsilon_{iv} \ , \quad \text{(C1)}$$

where $X_{iv}$ is the covariate, $I_{\{T_{iv}=Pk\}}$, k=2,3,4 are indicators for the treatments P2, P3 and P4 (P1 is the omitted category) and $S_v$ with $v \in \{1, \dots, 24\}$ represents strata dummies.[24] The error term $\varepsilon_{iv}$ is clustered at the village level. The test for joint significance of $\beta_2, \beta_3$ and $\beta_4$ is thus equivalent to a test for equal means in the treatment arms P1 to P4.

Table C1(a) provides summary statistics and balance tests for sociodemographic, economic and health indicators on the household and individual levels from the baseline survey. Comparing the means of sociodemographic indicators across treatment groups (first panel), we observe no significant differences. This is confirmed by the relatively high p-values of the joint test for model significance. The economic indicators (second panel), household health indicators (third panel) and individual health indicators (fourth panel) show no statistically significant differences across treatment groups. Table C1(b) provides summary statistics and balance tests for the bi-monthly phone survey data. Consent to the phone survey is above 90% and balanced across treatments.[25] About 2% of dependents report an inpatient event, leading to 15% of households having had some dependent admitted. These numbers are similar to the health-seeking behavior at baseline. Again, all variables appear to be balanced across treatment arms. Balancing also holds when the two control groups of villages are included where no additional insurance was available in the comparison.

---

[24] Note that strata fixed effects are included only in the balance tests for the main treatments P1 to P4. Discounts are randomized on the level of the household and thus not stratified.

[25] We conduct a separate attrition analysis, but do not find any systematic differences in drop-out across the treatments.

**Table C1-Balance Tests across Insurance Policy Treatments**

**(a) Baseline Survey**

| | Overall | P1 | P2 | P3 | P4 | P-val |
|---|---|---|---|---|---|---|
| *Socio-Demographics - HH* | | | | | | |
| HH Size (Survey) | 5.99 | 5.95 | 5.95 | 6.03 | 6.03 | 0.57 |
| | (2.117) | (2.093) | (2.072) | (2.054) | (2.237) | |
| HH Size (Matched) | 5.37 | 5.26 | 5.43 | 5.37 | 5.42 | 0.37 |
| | (1.912) | (1.872) | (1.956) | (1.822) | (1.986) | |
| Dependents (Matched) | 3.59 | 3.48 | 3.62 | 3.59 | 3.65 | 0.44 |
| | (1.869) | (1.834) | (1.876) | (1.791) | (1.961) | |
| Age of Client | 38.62 | 38.85 | 38.57 | 38.24 | 38.82 | 0.69 |
| | (10.887) | (10.918) | (10.934) | (10.741) | (10.955) | |
| Client Female | 0.53 | 0.57 | 0.51 | 0.50 | 0.54 | 0.33 |
| | (0.499) | (0.495) | (0.500) | (0.500) | (0.499) | |
| Client No Education (D) | 0.55 | 0.56 | 0.52 | 0.55 | 0.56 | 0.37 |
| | 5.99 | 5.95 | 5.95 | 6.03 | 6.03 | |
| *Economic - HH* | | | | | | |
| Income (month) | 22691 | 21623 | 24515 | 22627 | 21953 | 0.28 |
| | (24695) | (20018) | (34658) | (20225) | (20379) | |
| Asset Index | 0.06 | 0.06 | 0.20 | 0.07 | -0.09 | 0.37 |
| | (2.422) | (2.433) | (2.539) | (2.319) | (2.387) | |
| Savings | 12085 | 13548 | 13092 | 10147 | 11607 | 0.64 |
| | (67986) | (71670) | (85948) | (31357) | (70158) | |
| Credit | 30439 | 27603 | 33057 | 30112 | 30803 | 0.35 |
| | (71910) | (54074) | (79531) | (78197) | (72204) | |
| *Health & Insurance - HH* | | | | | | |
| Any Inpatient (D) | 0.12 | 0.11 | 0.13 | 0.12 | 0.12 | 0.51 |
| | (0.327) | (0.316) | (0.338) | (0.325) | (0.328) | |
| Knows Health Insurance (D) | 0.18 | 0.20 | 0.19 | 0.18 | 0.16 | 0.62 |
| | (0.385) | (0.397) | (0.390) | (0.383) | (0.369) | |
| *Health - Dependents* | | | | | | |
| Health Step (1-5) | 4.76 | 4.75 | 4.76 | 4.75 | 4.77 | 0.97 |
| | (0.631) | (0.631) | (0.644) | (0.648) | (0.602) | |
| Outpatient Experience (D) | 0.14 | 0.14 | 0.15 | 0.15 | 0.14 | 0.96 |
| | (0.351) | (0.349) | (0.355) | (0.353) | (0.346) | |
| Inpatient Experience (D) | 0.02 | 0.02 | 0.02 | 0.01 | 0.02 | 0.60 |
| | (0.126) | (0.124) | (0.135) | (0.121) | (0.124) | |
| Outpatient Cost | 609.99 | 302.63 | 706.62 | 491.23 | 895.49 | 0.00 |
| | (7920.430) | (2198.461) | (9268.894) | (5763.700) | (10873.455) | |
| Inpatient Cost | 506.36 | 404.38 | 747.68 | 429.66 | 434.93 | 0.40 |
| | (7520.868) | (5261.274) | (11433.696) | (6260.521) | (5164.277) | |
| N (Dependents) | 15361 | 3560 | 3920 | 3796 | 4085 | |
| N (HHs) | 4283 | 1022 | 1083 | 1058 | 1120 | |

**(b) Phone Survey**

| | Overall | P1 | P2 | P3 | P4 | P-val |
|---|---|---|---|---|---|---|
| Consent to participate (D) | 0.93 | 0.92 | 0.93 | 0.93 | 0.93 | 0.82 |
| | (0.259) | (0.269) | (0.254) | (0.261) | (0.253) | |
| *Health - HH* | | | | | | |
| Any Inpatient (D) | 0.14 | 0.15 | 0.13 | 0.15 | 0.15 | 0.46 |
| | (0.351) | (0.353) | (0.334) | (0.360) | (0.355) | |
| Any Outpatient (D) | 0.65 | 0.66 | 0.66 | 0.64 | 0.65 | 0.85 |
| | (0.476) | (0.475) | (0.473) | (0.480) | (0.478) | |
| *Health - Dependents* | | | | | | |
| Inpatient Experience (D) | 0.02 | 0.02 | 0.01 | 0.01 | 0.02 | 0.96 |
| | (0.124) | (0.130) | (0.121) | (0.120) | (0.124) | |
| Outpatient Experience (D) | 0.14 | 0.14 | 0.14 | 0.14 | 0.14 | 0.88 |
| | (0.348) | (0.348) | (0.349) | (0.350) | (0.344) | |
| Inpatient Cost | 371.59 | 438.46 | 452.54 | 371.85 | 237.36 | 0.12 |
| | (5537.914) | (5116.372) | (8022.091) | (4937.016) | (2872.399) | |
| Outpatient Cost | 702.79 | 569.42 | 769.31 | 638.28 | 812.38 | 0.07 |
| | (5415.117) | (3154.431) | (5475.952) | (5350.682) | (6772.168) | |
| N (Dependents) | 14246 | 3271 | 3641 | 3496 | 3834 | |
| N (HHs) | 4283 | 1022 | 1083 | 1058 | 1120 | |

*Notes:* The table provides means and standard deviations (in parentheses) of the variables. Column 1 provides overall measures, while other columns indicate the policy. The last column contains the p-value from a joint test for model significance of equation (C1) including strata fixed effects. Standard errors are clustered at the village level. Binary variables are indicated with (D).

In a next step, we provide evidence for a balanced distribution of discount vouchers. Random assignment through household-level lotteries with replacement implies an expected uniform probability distribution of discounts. Table C2 illustrates the frequencies of the four discount levels across insurance policy as well as in general. In addition, we test the null-hypothesis of the expected uniform distribution by Pearson's Chi-square test, the p-value of which is reported in the second to last row. Our test does not reject the hypothesis of a uniform distribution, even though the share of zero discounts is lower than 25%. This holds true also for policy P1 for which we observe a stronger deviation from the expected uniform distribution.

**Table C2 - Balance Check: Discount Allocation**

| | P1 | P2 | P3 | P4 | Overall |
|---|---|---|---|---|---|
| 0 | 0.19 | 0.23 | 0.22 | 0.22 | 0.22 |
| 10 | 0.27 | 0.27 | 0.26 | 0.28 | 0.27 |
| 20 | 0.27 | 0.28 | 0.25 | 0.27 | 0.27 |
| 30 | 0.27 | 0.23 | 0.27 | 0.23 | 0.25 |
| Pearson Chi2 P | 0.2268 | 0.4632 | 0.5998 | 0.2290 | 0.2144 |
| HHs | 856 | 870 | 830 | 877 | 3432 |

*Notes:* Relative frequencies of discounts given the respective policy. Pearson Chi2 p provides the p-value from a chi-square test with H0 of a uniform distribution. The difference in number of observations to the main balance checks is explained by the fact that only households attending the community meeting received a discount.

To investigate potential systematic imbalances, we provide additional tests in Table C3. The idea is to investigate whether specific household characteristics, potentially related to health indicators and thus insurance demand, cause a jump in the probability of receiving a specific discount voucher. We replace the main treatment indicators in equation (C1) with discount-level indicators, where the zero discount group serves as the reference group. We test for discontinuous jumps in the probability of receiving a specific discount by conducting a joint test for model significance. The final column provides the corresponding p-value. We observe no statistically significant difference across discount levels for any of the health indicators. Similarly, there are no systematic differences in economic indicators. In terms of socio-demographic variables, it seems that there are statistically significant differences in the age and sex composition across discount levels. A clear, systematic pattern such as older individuals or females receiving higher discounts, however, is not visible. For this reason, we are confident that the randomization of discounts through household lotteries in the field is not subject to systematic imbalances.

**Table C3 - Balance Checks (Discounts)**

| | Overall | D=0 | D=10 | D=20 | D=30 | P-val |
|---|---|---|---|---|---|---|
| *Socio-Demographics – HH* | | | | | | |
| HH Size | 5.99 | 5.98 | 5.96 | 6.01 | 6.01 | 0.94 |
| | (2.10) | (2.03) | (2.05) | (2.24) | (2.08) | |
| Age of Client | 38.72 | 38.33 | 39.52 | 39.02 | 37.86 | 0.01 |
| | (10.96) | (10.92) | (11.22) | (11.19) | (10.40) | |
| Client Female (D) | 0.53 | 0.50 | 0.52 | 0.57 | 0.54 | 0.03 |
| | (0.50) | (0.50) | (0.50) | (0.50) | (0.50) | |
| Client No Education (D) | 0.54 | 0.53 | 0.55 | 0.57 | 0.52 | 0.16 |
| | (0.50) | (0.50) | (0.50) | (0.50) | (0.50) | |
| *Economic – HH* | | | | | | |
| Avg. Inc. (month) | 22727.40 | 22963.60 | 21587.96 | 24125.41 | 22264.71 | 0.12 |
| | (25552.58) | (30839.78) | (16445.10) | (28186.03) | (25640.41) | |
| Land (acres) | 1.41 | 1.29 | 1.48 | 1.42 | 1.45 | 0.66 |
| | (3.26) | (2.92) | (3.29) | (3.12) | (3.65) | |
| Savings | 12343.92 | 9757.70 | 14193.15 | 12840.66 | 12043.18 | 0.40 |
| | (73131.30) | (33068.48) | (85167.08) | (90299.23) | (62995.56) | |
| Credit | 30861.74 | 30574.92 | 32890.37 | 30272.85 | 29535.92 | 0.72 |
| | (70147.88) | (80249.15) | (65293.04) | (73655.31) | (61564.62) | |
| *Health & Insurance* – HH | | | | | | |
| Any Inpatient (D) | 0.12 | 0.13 | 0.13 | 0.11 | 0.12 | 0.55 |
| | (0.33) | (0.34) | (0.33) | (0.31) | (0.32) | |
| Total Inpatient Cost | 4379.73 | 4895.13 | 4972.60 | 3658.71 | 4060.64 | 0.55 |
| | (22281.90) | (24974.56) | (26317.16) | (19502.17) | (17260.48) | |
| Knows Health Ins. (D) | 0.18 | 0.18 | 0.21 | 0.17 | 0.16 | 0.07 |
| | (0.39) | (0.38) | (0.41) | (0.38) | (0.37) | |
| N (Dependents) | 12283 | 2643 | 3283 | 3236 | 3121 | |
| N (HHs) | 3433 | 740 | 927 | 913 | 853 | |

*Notes:* The table provides means and standard deviations (in parentheses) of the respective variables. Column provides overall measures, while other columns indicate the respective policy. The last column contains the p-value from a joint test for model significance of equation (C1). Standard errors are clustered at the village level. Binary variables are indicated with (D).

Table C5 provides analogous balance tests for the group meeting attendance. We observe no statistically significant differences in observables between meeting attendants and non-attendants. The observed similarity supports external validity of our results for the population of credit clients in Sargodha district.

**Table C5 - Balance Checks (Meeting Attendance)**

| | Overall | Not Attending | Attending | P-val |
|---|---|---|---|---|
| *Health - Dependent* | | | | |
| Expected Reimbursement Cost (PKR)^ | 82.73 | 82.59 | 82.77 | 0.95 |
| | (134.352) | (138.970) | (133.176) | |
| Subjective Health Status (1-5) | 4.76 | 4.77 | 4.76 | 0.41 |
| | (0.631) | (0.625) | (0.633) | |
| Outpatient Treatment (D) | 0.14 | 0.14 | 0.14 | 0.69 |
| | (0.351) | (0.348) | (0.351) | |
| Inpatient Treatment (D) | 0.02 | 0.02 | 0.02 | 0.88 |
| | (0.126) | (0.125) | (0.127) | |
| Outpatient Cost (PKR) | 609.99 | 417.14 | 658.32 | 0.06 |
| | (7920.430) | (5189.928) | (8467.294) | |
| Inpatient Cost (PKR) | 506.36 | 525.34 | 501.60 | 0.85 |
| | (7520.868) | (6632.404) | (7727.760) | |
| *Socio-Demographics - HH* | | | | |
| HH Size (Survey) | 5.99 | 5.99 | 5.99 | 0.97 |
| | (2.117) | (2.170) | (2.104) | |
| Age of Client | 38.62 | 38.23 | 38.72 | 0.23 |
| | (10.887) | (10.596) | (10.958) | |
| Client Female (D) | 0.53 | 0.52 | 0.53 | 0.74 |
| | (0.499) | (0.500) | (0.499) | |
| Client Has No Education (D) | 0.55 | 0.56 | 0.54 | 0.34 |
| | (0.498) | (0.497) | (0.498) | |
| *Economic - HH* | | | | |
| Avg. Monthly Earning (PKR) | 22691.34 | 22560.66 | 22723.69 | 0.86 |
| | (24694.608) | (20900.437) | (25549.780) | |
| Asset Index | 0.06 | -0.05 | 0.09 | 0.13 |
| | (2.422) | (2.419) | (2.423) | |
| Savings (PKR) | 12085.10 | 11054.26 | 12340.33 | 0.48 |
| | (67986.387) | (41200.112) | (73120.945) | |
| Total Credit (PKR) | 30438.72 | 28731.23 | 30861.49 | 0.41 |
| | (71910.002) | (78684.167) | (70137.665) | |
| *Health & Insurance - HH* | | | | |
| Any Inpatient (D) | 0.12 | 0.12 | 0.12 | 0.87 |
| | (0.327) | (0.325) | (0.327) | |
| Inpatient Cost (HH) | 4445.89 | 4718.24 | 4378.45 | 0.76 |
| | (24474.931) | (31854.398) | (22278.776) | |
| Knows Insurance (D) | 0.18 | 0.18 | 0.18 | 0.73 |
| | (0.385) | (0.381) | (0.386) | |
| N (Dependents) | 15361 | 3078 | 12283 | |
| N (HHs) | 4283 | 850 | 3433 | |

*Notes:* The table provides means and standard deviations (in parentheses) of the respective variables. Column 1 provides overall measures, while other columns indicate the attendance of the respective household in the group meeting. The last column contains the p-value from a joint test for model significance similar to equation (C1), excluding strata fixed effects. Standard errors are clustered at the village level. Binary variables are indicated with (D). Monetary variables are in Pakistani Rupees (PKR).

^ In line with the other balancing tables, we include all treatment arms in the test – including the individual high insurance (P2), which features higher expected costs. The mean of the costs index is therefore somewhat higher than in the standard coverage treatments only (P1, P3, P4).

## D. Construction of the Health Risk Index

The insurer's average cost curve constitutes the central element for testing adverse selection in this study. A straightforward estimate of the average cost curve would aggregate the insurer's reimbursed claims for a given insurance product and price level.[26] Since hospitalization is a rare event, we cannot – despite the large sample size – directly estimate the average cost curve based on these reimbursed claims. Instead, we use detailed baseline health and demographic information ($X_{i0}$) to predict the insurance provider's reimbursement costs for each individual *i* ($C_{i1}$). Time is indicated with 0 at baseline and with 1 at the end of the insurance period. We are interested in obtaining a good estimate of the conditional expectation of the provider's reimbursement cost at the end of the insurance period: $\hat{E}[C_{i1}|X_{i0}]$.

Again, a direct approach would use the observed reimbursement cost to estimate their relation to baseline characteristics. However, claims are too rare in our data to obtain a good estimate (only 39 claim cases are reported). This is partly because claims can only be made by people who are insured, which excludes the non-insured part of our sample from such an analysis. Furthermore, not all hospitalization cases lead to a claim.[27] We therefore use detailed information on inpatient health events and costs, gathered in our bi-monthly phone survey during the one-year product cycle. Hospitalization events in the phone survey are reported for 334 of the 21,470 dependents in the phone survey sample. Based on the aggregated inpatient expenditures during the insurance period, we calculate the maximum amount for each individual that could be reimbursed under the insurance policy ($\bar{C}_{i1}$). Subsequently, $\hat{E}[\bar{C}_{i1}|X_{i0}]$ can be predicted using an adequate regression. We account for the fact that not all of these costs are claimed by adjusting the final expected cost index ($ECI_{i1}$) as follows:

$$ECI_{i1} = \hat{E}[\bar{C}_{i1}|X_{i0}] \times \frac{\sum_P \sum_{i \in Insured} C_{iP1}}{\sum_P \sum_{i \in Insured} \hat{E}[\bar{C}_{i1}|X_{i0}]} \qquad \text{(D1)}$$

This means the prediction is made based on all potentially claimable costs, which maximizes statistical power. At the same time, the index is scaled by the ratio between actual claim amounts

[26] As described in section III, there are four insurance products and four price levels.

[27] To gain insights into this phenomenon we conducted in-depth interviews with some households that were insured, reported a hospitalization event and yet did not claim reimbursement of their expenses. These interviews were conducted after the end of the insurance period to avoid interfering with the research study. The reported reasons for this behavior are manifold. While some incidences can be explained with unawareness about the claim procedure or frustration about the process, other cases are related to missing trust, preference for alternative (more expensive) coping strategies and recall problems about having bought the insurance product.

relative to the maximal claimable amount according to the policy (PKR 15,000 for P1, P3, and P4).

We estimate $\hat{E}[\bar{C}_{i1}|X_{i0}]$ using a Tobit model, controlling for a broad range of baseline household- and individual-level characteristics.[28] The household-level variables account for the economic situation, the household size and client characteristics. The individual-level characteristics include demographic information such as age, gender and whether the individual is contributing to the household income as well as detailed health information. The latter includes individuals' subjective health status, inpatient and outpatient health history, associated costs, type of health events experienced and subjective magnitude of the shocks experienced. Table D1 reports the estimated coefficients in the Tobit regression for eligible dependents. The estimated coefficients suggest that dependents in younger age groups cause lower reimbursable claims than the reference group of 30- to 49-year-olds. Further, better subjective health and better self-reported health history result in lower reimbursable costs.

Column 2 of Table D1 reports the results of an identical estimation that considers only the eligible dependents in the control groups. The purpose of this additional regression is to assess the robustness of our results to the existence of moral hazard. As described in Section III, the control groups are not offered any additional insurance and hospitalization expenditure for dependents in this group and hence should not be affected. Thus, comparing the coefficient estimates in columns 1 and 2 shows whether moral hazard changes the mapping from baseline characteristics to hospitalization expenditures. The resulting coefficient estimates are mostly similar to the ones reported in column 1 in sign and magnitude. Based on a Hausman specification test, we cannot reject that both models are equivalent (p-value: 0.57). This is consistent with the fact that there is no significant treatment effect of the insurance treatments on inpatient expenditures (see Table D3). The choice between including all observations and using the control groups only therefore does not make a large difference. To maximize the precision of our estimates, we include all observations (i.e. specification 1).

[28] A Tobit model is a natural choice, as maximum claimable amounts cannot be lower than zero and are restricted to PKR 15,000 in policies P1, P3 and P4.

**Table D1 - Predicting Inpatient Expenditure using Baseline Characteristics**

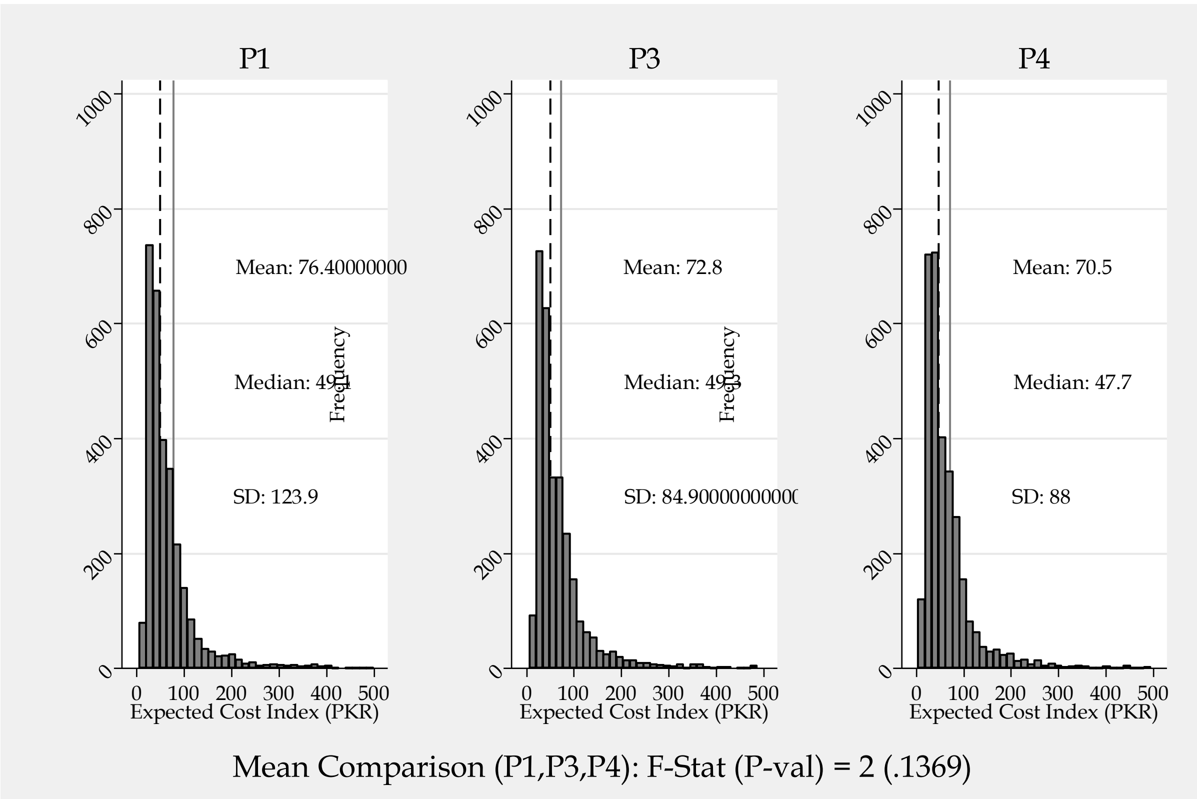
| | 1 | 2 |
|---|---|---|
| | All T | Controls only |
| *Household Level Info* | | |
| HH Size | -2117.81*** | -2157.93* |
| | (616.24) | (1157.82) |
| Income (in 1000 Rs.) | 56.33 | -0.41 |
| | (43.62) | (50.86) |
| Saving (in 1000 Rs.) | 15.27 | 20.89 |
| | (10.48) | (22.90) |
| Asset Index | 178.60 | -565.30 |
| | (519.33) | (748.02) |
| Client Female | -2693.87 | -3971.38 |
| | (2501.32) | (3516.04) |
| Client has no education | -225.60 | -1862.40 |
| | (2464.48) | (3787.29) |
| *Individual Level Info* | | |
| Age (0-4) | -11284.12** | -2515.22 |
| | (5128.38) | (7817.78) |
| Age (5-9) | -23400.18*** | -16939.28** |
| | (5535.32) | (8145.62) |
| Age (10-14) | -25454.44*** | -17694.00** |
| | (5849.77) | (8555.38) |
| Age (15-19) | -12717.43*** | -8448.89 |
| | (4826.51) | (7032.81) |
| Age (20-29) | -8764.89* | -9052.73 |
| | (5133.15) | (7650.28) |
| Age (50-59) | 1512.15 | -539.27 |
| | (6570.36) | (10119.70) |
| Age (60-69) | -5043.40 | -5590.26 |
| | (6590.69) | (9704.61) |
| Age (70+) | -4342.60 | -478.83 |
| | (7169.47) | (10114.03) |
| Working | -14635.29*** | -16080.74** |
| | (4194.80) | (6587.80) |
| Female | 266.71 | -1961.23 |
| | (2447.19) | (3387.72) |
| Subjective Health Status (1-5) | -6667.32*** | -7474.08*** |
| | (1659.26) | (2535.78) |
| Outpatient Treatment | 8348.97* | 90.09 |
| | (4355.64) | (6834.38) |
| Inpatient Cost (PKR) | 0.06 | 0.02 |
| | (0.07) | (0.07) |
| Outpatient Cost (PKR) | 0.07*** | 0.06*** |
| | (0.02) | (0.01) |
| Chronic Inpatient Disease | 31643.95*** | 12698.73 |
| | (9441.58) | (12701.57) |
| # Inpatient Cases | 2157.61*** | 7044.08* |
| | (826.55) | (4194.29) |
| # Neglected Inpatient Care | -1061.03 | -967.46 |
| | (2798.62) | (3883.62) |
| Drop in Subj. Health (Inpatient) | -5246.08** | -3710.95 |
| | (2505.71) | (4309.28) |
| Drop in Subj. Health (Outpatient) | 339.89 | -1550.96 |
| | (1624.50) | (2371.66) |
| Constant | -48526.07*** | -30326.65* |
| | (10454.33) | (15573.63) |
| sigma | 49197.76*** | 42658.61*** |
| | (3485.41) | (4775.13) |
| N | 21473 | 7227 |
| F value | 6.30 | 4.60 |
| P-value of F statistic | 0.0000 | 0.0000 |

*Notes:* The table provides results from a Tobit model that explains the maximal claimable costs as a function of household- and individual-level variables. Standard error in parentheses are clustered at the village level. Monetary amounts are in Pakistani rupees (PKR), where 101 PKR ≈ USD 1.

We predict expected claimable inpatient expenditures $\hat{E}[\bar{C}_{i1}|X_{i0}]$ for each individual using specification 1 of Table D1. Consistent with Equation D1, we then apply a scaling factor of 0.4588 to predict the expected cost index $ECI_{i1}$ for each individual under the policy.[29]

Figure D1 illustrates the distribution of the expected insurer costs across policies P1, P3, and P4. The mean of the distribution is shown as a grey solid and the median as a black dashed line. The figure reveals that the cost distribution is right-skewed in a similar way for all policies. A test for equality of their means cannot be rejected (p-value: 0.1369).

**Figure D1 – Histogram of the expected cost index, by policy**



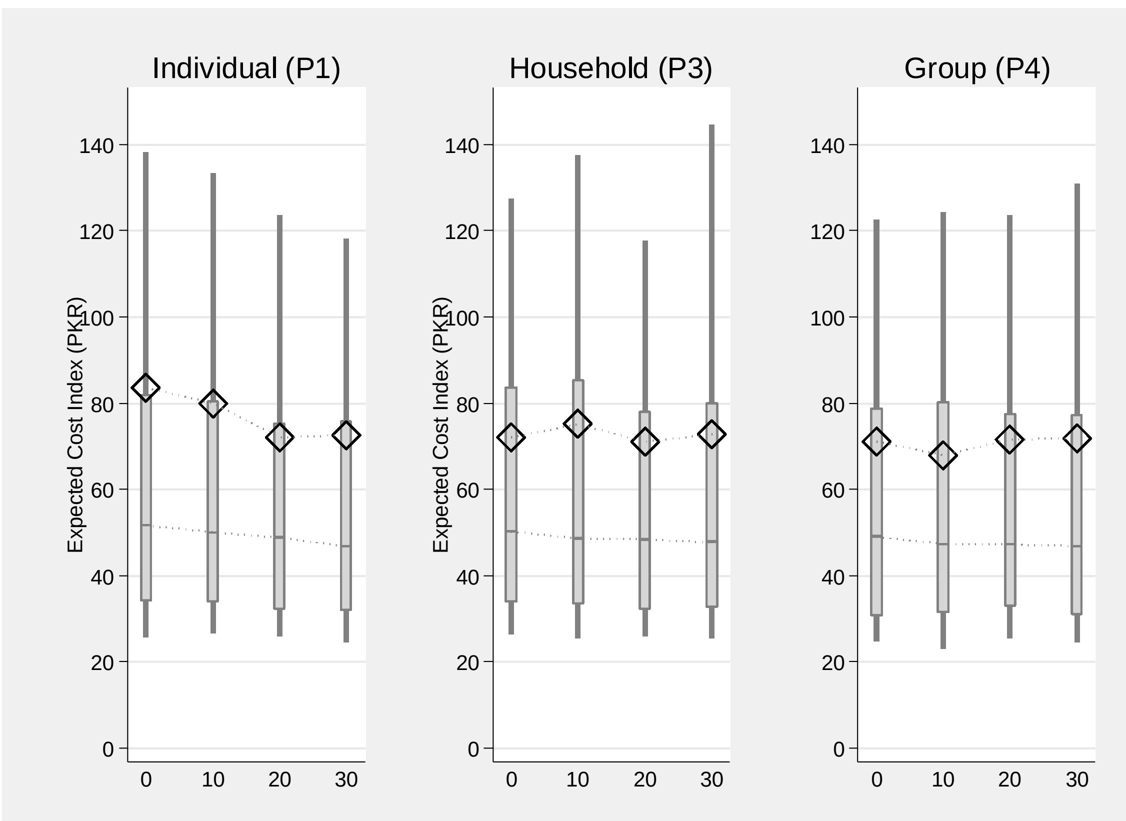

*Notes:* The figures shows histograms of the provider's expected reimbursement costs across the four policies. The mean and median are illustrated through the solid and dashed line respectively. The predicted reimbursement costs are measured in Pakistani rupees (PKR), where 101 PKR ≈ USD 1.

Figure D2 shows the balancing of the cost index across policies and prices. The box plots illustrate the interquartile rage (IQR), in addition to the 10th and the 90th percentile of the distribution. The distributions appear balanced across prices in all policies.

[29] The scaling factor is based on hospitalization expenditure and claim data during the insurance period as summarized in Table D2.

Table D2 summarizes and compares hospitalization costs up to the theoretical coverage limit ("Claimable Inpatient Costs"), number of claims reimbursed and average payouts under the different insurance policies. Reimbursed claims are based on all observations in the insurance data set. Claimable costs are based on the self-reported information from the bi-monthly phone survey and restricted to the observations that can be matched with insurance data (the dataset used in the paper). Matched and non-matched observations from the survey data are not significantly different, though. Besides illustrating the ratio between insurance payouts and potentially claimable amounts (0.3885), the table reveals that there are indeed strong differences in paid claims between products. The payout frequency tends to be higher in individual policies (P1, P2) than in households or group policies (P3, P4) and despite the limited number of cases, several comparisons via two-sample proportion tests are significant: P1 vs. P4 (p-value: 0.0782), P2 vs. P3 (p-value: 0.0216) , P2 vs. P4 (p-value: 0.0133) and P1+P2 vs. P3+P4 (p-value: 0.0054). Comparisons of P1 vs. P2 and P3 vs. P4 are all insignificant.

**Figure D2 – Distribution of risk across discounts and policy regimes**




*Notes*: This figure illustrates the distribution of the expected cost index by discount level and policy regime. The box plot illustrates the interquartile range (IQR), with the median indicated by the line separating the box. The lower (upper) adjacent line shows the 90th (10th) percentile, respectively. The diamond indicates the value of the mean.

**Table D2 - Summary Statistics of Inpatient Expenditure and Claim Behavior**

| | N Insured | N Insured (Matched) | Mean Claimable Inpatient Costs^ | Mean Predicted Claimable Inpatient Costs^ | N Claims (Total)~ | Mean Amount Claimed~ |
|---|---|---|---|---|---|---|
| P1 | 1054 | 921 | 349.59 | 212.11 | 12 | 114.18 |
| P2 | 663 | 615 | 450.90 | 316.72 | 11 | 202.36 |
| P3 | 1505 | 1350 | 166.80 | 169.62 | 9 | 59.21 |
| P4 | 1344 | 1212 | 122.69 | 163.10 | 7 | 55.04 |
| Total | 4566 | 4098 | 235.55 | 199.36 | 39 | 91.46 |

*Notes:* Monetary amounts are in Pakistani rupees (PKR), where 101 PKR ≈ USD 1. “Insured” are all individuals appearing the insurance management information system, “Insured (Matched)” are those Insured that can be matched with our survey data. ^ Based on “Insured (Matched)”, ~ based on “Insured”.

**Table D3 – Treatment Effect of Insurance Policies on Reported Inpatient Cost**

| | Inpatient Cost (PKR) |
|---|---|
| P1 | 158.3321* |
| | (92.7487) |
| P3 | 109.9905 |
| | (106.1380) |
| P4 | -44.5723 |
| | (62.1140) |
| Strata FE | yes |
| N | 17832 |
| $R^2$ | 0.0014 |
| Wald | 1.6900 |
| p(Wald) | 0.1685 |

*Notes:* Reported inpatient costs are in Pakistani rupees (PKR), where 101 PKR ≈ USD 1. The control group serves as the reference group. The OLS regression includes strata fixed effects and standard errors are clustered at the village level. The Wald test statistic is from a joint test of significance of the main treatment indicators. The estimation sample contains eligible dependents of all policies, excluding policy P2, for which there is information from the follow-up phone survey.