American Economic Journal: Applied Economics 2023, 15(3): 313–340
https://doi.org/10.1257/app.20200639

Adverse Selection in Low-Income Health Insurance Markets: Evidence from an RCT in Pakistan†

By Torben Fischer, Markus Frölich, and Andreas Landmann*

We present robust evidence on adverse selection in hospitalization insurance for low-income individuals that received first-time access to insurance. A large randomized control trial from Pakistan allows us to separate adverse selection from moral hazard, estimate how selection changes at different points of the demand curve, and test simple measures to limit 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 level or higher almost eliminates adverse selection, thus mitigating its welfare consequences and facilitating sustainable insurance supply. (*JEL* D82, G22, I13, I18, O15, O16)

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, they 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

* Fischer: IDinsight (email: torben.fischer@idinsight.org); Frölich: Department of Economics, University of Mannheim and Center for Evaluation and Development (C4ED) (email: froelich@uni-mannheim.de); Landmannc: Institute of Economics, Friedrich-Alexander-Universität Erlangen-Nürnberg (email: andreas.landmann@fau.de). Neale Mahoney was coeditor for this article. 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, Amy Finkelstein, 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 First Evidence Based Economics Summer Meeting, the EDePo Workshop, the 2016 TAPES Conference on Social Insurance, the MACIE Research Seminar at the University of Marburg, the Development Seminar at the Paris School of Economics, the Twelfth 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 Seventieth 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, Fawad Ejaz, Teresa Leikard, and Sarah Syring. 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.

† Go to https://doi.org/10.1257/app.20200639 to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

market-based solutions is large, despite the de facto inexistence of insurance markets in many cases.[1] Whether insurance schemes can enter such "missing markets" and attain efficiency critically 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 2017; Banerjee et al. 2019), 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, struggle to separately identify 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 (a notable exeption is Yao, Schmit, and Sydnor 2017). Second, almost none of these research settings allows a rigorous assessment of the welfare consequences of adverse selection. Third, there is little 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 and de facto provide first-time access to health insurance for all household members. 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 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. One limitation of the study is that we only observe 39 claims, as (i) hospitalization is a rare event, (ii) only about 21 percent of our observations end up in the insurance pool, and (iii) some reported hospitalization events do not lead to a claim (which we cannot fully explain). To construct cost curves with sufficient power, we therefore estimate the insurers' expected reimbursement costs for each individual's inpatient expenditures based on all 334 inpatient events reported

[1] The Swiss Reinsurance Company estimates that the microinsurance (i.e., low-income insurance) market comprises approximately four 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).

in our high-frequency phone survey. To predict expected reimbursement costs, we use detailed baseline health status, health history, and many other baseline characteristics (more details in Section V and online Appendix D).

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 and leading to the threat of a market breakdown. When bundling insurance policies at the household or group level, however, adverse selection is mitigated. A welfare analysis suggests that bundled policies can sustain substantially higher quantities and lower prices than individual policies. Further, the welfare consequences of adverse selection seem less severe in relative terms for household policies. Our results highlight the important role of adverse selection for the existence of a market. Such evidence on whether policies can be offered at all is rare (Cabral 2017), particularly in the context of a nascent insurance market.

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

I. 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 those for the noninsured. 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 (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., Cameron et al. 1988; 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, it 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

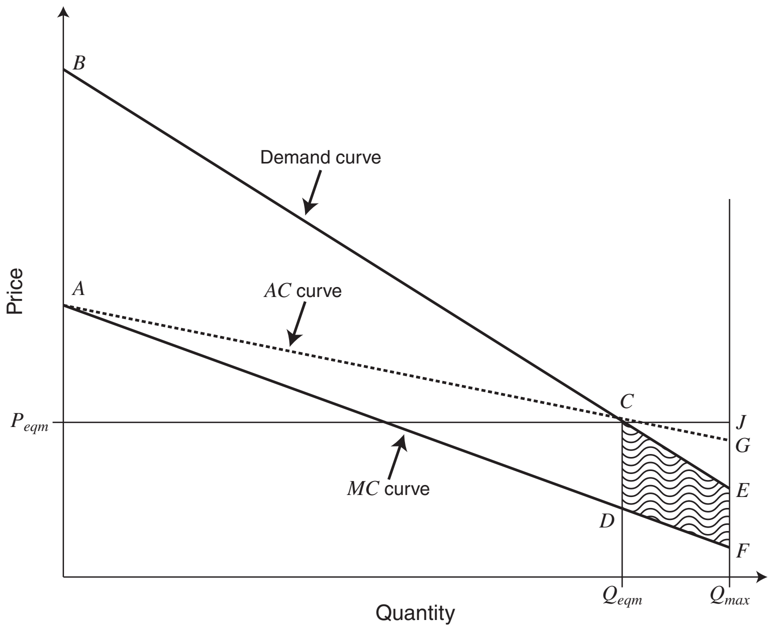

Figure 1. Analysis of Adverse Selection and Welfare

Source: Figure 1 from Einav and Finkelstein (2011)

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 upward shift of the average cost curve still reflects the degree of adverse selection.[3] De Meza and Webb (2017) provide a more detailed discussion in favor of exploiting price variation along similar lines.

A prerequisite 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 this identification strategy. They investigate the presence of adverse selection and its implied welfare costs in the context of employer-provided health insurance in the United States. Using countrywide data from a large US employer, they exploit

[2] The finding of advantageous selection would not be in line with classical adverse selection model but could result, for example, 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 upward 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 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.

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 quasi-experimental identification approach (e.g., Hackmann, Kolstad, and Kowalski 2012, 2015; Panhans 2019; Finkelstein, Hendren, and Shepard 2019), we only know of one recent study using experimental variation in premium prices to assess adverse selection (Banerjee et al. 2019, in the context of Indonesia).

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. These cost data are 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 III, discuss the "expected cost index" in Section V, and provide more details on its construction in online Appendix D.

II. 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 Section IIA. The second subsection explains the interventions as well as the most important hypotheses linked to each policy. Section IIC presents our sampling strategy and the randomization procedures used for treatment allocation.

A. *Background*

According to World Bank data from 2015, Pakistan is a lower-middle-income country with a population of almost 200 million and a GDP per capita of US$1,357. About one-quarter of the population lives below the national poverty line.[5] Furthermore, most households are at risk of remaining or falling into poverty (World Bank 2007, 2008). The government spends less than 1 percent of its GDP on health, which is low even for a developing country. Public health expenditure hence accounts for only 28 percent of total health expenditure; 69 percent of health expenditure is private, almost all of which has to be paid out of pocket. 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 at http://data.worldbank.org/country/pakistan. Subsequent figures on public health spending and out-of-pocket expenditures from 2015 (the time of our intervention) 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. Very few nongovernmental organizations (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 2011, 2012).[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 microenterprise development loans to smaller, jointly liable credit groups that usually consist of three to six members.

Since 2005, NRSP has complemented its microcredit 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 15,000 Pakistani rupees (PKR) (about US$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 toward funeral charges is paid to the family in the case of death (except suicide) 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 transportation costs. For maternity expenses, the reimbursement threshold is set to PKR 10,000. Preexisting 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

[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 (https://tribune.com.pk/story/1010674/sehat-sahulat-programme-k-p-to-get-social-health-protection).

[7] The maximal benefit depends on the degree of disability caused by the accident.

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.

B. *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 grassroots 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 nonpanel 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.[9] 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 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, and the premium remains at PKR 100 per person insured. The *Group* policy (P4) requires at least 50 percent uptake within the credit group or CO. 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 for PKR 100 per individual. 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.[10] Note that in contrast to all other schemes, the high-coverage policy changes the expected reimbursement costs

[8] 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).

[9] Analyzing claims from the mandatory insurance in 2014 suggested a claim rate of about 1 percent and a fair premium of approximately PKR 100. This corresponds to 2–3 percent of average monthly income per capita in the sample.

[10] About 80 percent 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.

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; US\$1 ≈ PKR 101; PKR 15,000 ≈ US\$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 PKR 0, 10, 20, or 30 (pp) were randomized with equal probability at the household level.

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. Afterward, social organizers introduce the policy that has been randomly assigned to the community. During the sign-up phase, they privately offer each client a discount voucher for PKR 0, 10, 20, or 30, applicable to the per person premium for all eligible household members.

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 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 IV, V, and VI.

C. *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 distance and limited interaction between villages, this choice reduces 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 online 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 (MIS). 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.[11] 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.[12] 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 6,463 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.[13] In our analysis, we focus on the 334 villages in which the 4 insurance innovations have been implemented, with policies P1, P3, and P4 being of particular interest.

[11] In general, this translates into sampling one complete CO per village, sometimes amended by a smaller credit group. Alternatively, we sample four to five smaller credit groups per village. Note that according to the last census (1998), each village on average had a population of about 3,000 and 88 active loans before the start of the intervention.

[12] More details on the randomization procedure can be found in online Appendix B.

[13] This session is also conducted in the treatment villages in which an additional insurance policy is offered.

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	1,050	1,563
HHs	1,154	1,026	1,022	1,083	1,058	1,120	4,283	6,463
HHs attending	0	822	856	870	830	877	3,433	4,255
Dependents (dep.)	4,183	3,539	3,560	3,920	3,797	4,085	15,362	23,084
Attending dep.	0	2,798	2,981	3,209	2,938	3,156	12,284	15,082

Notes: Both the "Control" and "Awareness" groups are not used in this paper, except for the discussion of expected costs and moral hazard in online Appendix D.

As expected, the number of villages across treatment arms is balanced, and each treatment cluster comprises an average of 13 households.

III. Data

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

A. *Data Sources*

In the analysis, we combine household- and individual-level data from three sources. First, we use client-level information (NRSP 2016) captured in our implementation partner's 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 latter two combined constitute our household survey data (Fischer, Frölich, and Landmann 2019).

The MIS data include 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 the insurance eligibility of the dependents at the time of the insurance offer.[14] 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 versus accidental death/disability), the claim amount, and details on the disease diagnosed.

The household survey consists of several modules capturing sociodemographic, psychological, economic, and health indicators. The health module contains individual-level information on subjective health status, history of both in- and out-patient treatments, and detailed information on coping strategies. Baseline data were

[14] All household members in the roster, except for the microfinance client and his or her spouse (who have mandatory insurance), are defined as dependents. This procedure also ensures that the household structure is exogenous to the introduction of insurance.

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 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).

B. *Summary Statistics*

Table 3 shows some summary statistics for the 4,283 households in the 4 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 percent of the clients are female. Most clients have no formal education. The second panel of Table 3, panel A contains economic indicators. Average monthly income of households is about PKR 22,700 (US$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 PKR 30,000 and PKR 12,000, respectively. The third panel contains household-level health indicators. In about 12 percent of the sampled households, at least 1 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 18,000 per household. On average, 18 percent of the sampled households have heard about insurance. Sixteen percent of the dependents in the household consulted a doctor in the last month; 2 percent of household members had been hospitalized in the past 12 months. Panel B of Table 3 describes data gathered via the phone survey (93 percent of respondents in the baseline agree to be contacted via phone). During the 12 months covered, 15 percent 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 percent and 16 percent, respectively).

Online 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 percent, consistent with our uniform distribution scheme. Levels of discounts, furthermore, do not seem to differ by recipient characteristics.

Table 3—Summary Statistics

	Observations	Mean	SD
Panel A. Baseline characteristics			
Sociodemographics—HH			
HH size (survey)	4,283	5.99	2.12
HH size (matched)	4,283	5.37	1.91
Dependents (matched)	4,283	3.59	1.87
Age of client	4,283	38.62	10.89
Client female (D)	4,283	0.53	
Client no education (D)	4,283	0.55	
Economic—HH			
Income (month)	4,283	22,691	24,695
Asset index	4,283	0.06	2.42
Savings	4,283	12,085	67,986
Credit	4,283	30,439	71,910
Health and insurance—HH			
Any inpatient (D)	4,283	0.12	
Total inpatient cost	4,283	2,167	10,155
Knows health insurance (D)	4,283	0.18	
Health—dependents			
Health step (1–5)	15,361	4.76	0.63
Outpatient experience (D)	15,361	0.14	
Inpatient experience (D)	15,361	0.02	
Outpatient cost	15,361	212.05	878.73
Inpatient cost	15,361	383.65	4,181.49
Panel B. Phone survey			
Consent (D)	4,283	0.93	
Health—HH			
Any inpatient (D)	4,283	0.14	
Any outpatient (D)	4,283	0.65	
Health—dependents			
Inpatient experience (D)	14,246	0.02	
Outpatient experience (D)	14,246	0.14	
Inpatient cost	14,246	371.59	5,537.91
Outpatient cost	14,246	702.79	5,415.12

Notes: The table provides means and standard deviations (SD) of the respective variables. Binary variables are indicated with (D). Monetary amounts (rounded to integers if more than three digits) are in PKR, where PKR 101 $\approx$ US\$1. Positive baseline health costs (outpatient and inpatient) are winsorized at the ninetieth percentile.

IV. 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.[15] All of the offered policies uptake decreases in the premium.[16] The fraction of households covering some of their members is high in the individual policy (P1: 42–77 percent) compared to the household policy (P3: 26–74 percent) and the group policy (P4: 28–72 percent). In

[15] The figure is based on households attending the group meeting. Overall, around 80 percent 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 C4 in online Appendix C).

[16] Table A1 in online 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.

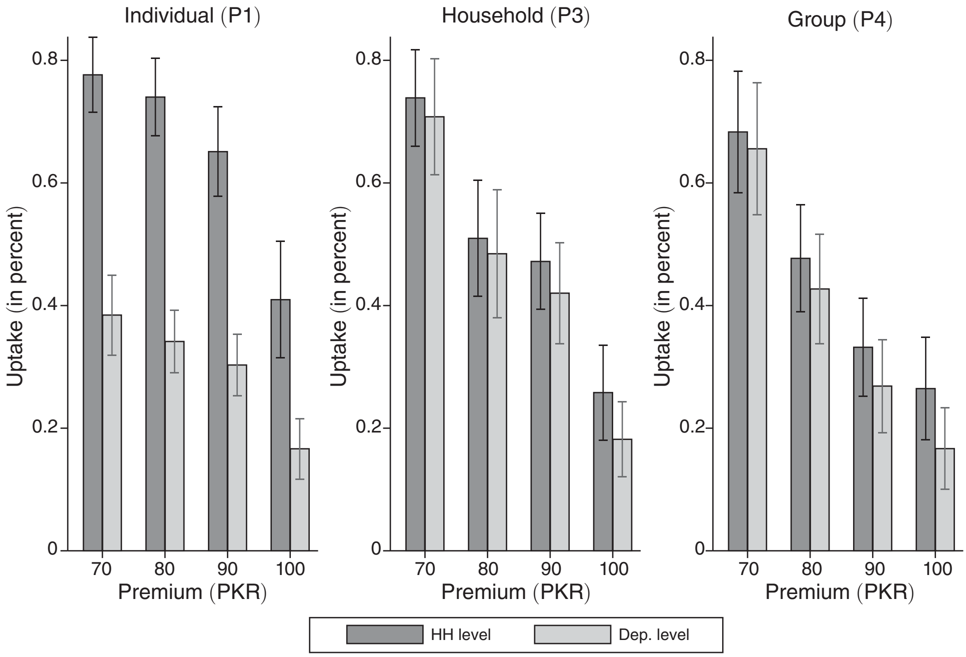

Figure 2. Insurance Demand, by Product Type

Notes: The bars indicate average uptake ratios on the household and dependent levels. The depicted 95 percent confidence intervals account for clustered standard errors at the village level. Small differences between dependent- and household-level uptake in policies P3 and P4 occur because of the smaller size of insured households. A more detailed breakdown is provided in Table 4.

terms of the fraction of dependents covered, however, the bundled policies achieve higher uptake (P3: 18–71 percent, P4: 19–68 percent) than the individual policy (P1: 17–39 percent). Table 4 shows the corresponding mean take-up numbers, standard errors, and number of insured for each policy-premium combination. It also includes formal test confirming differences between the individual, household, and group policies.[17] 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.

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 noninsured 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

[17] We reject the joint null hypotheses for the equality take-up on both the individual and household levels between policies P1 and P3 and policies P1 and P4 at different discount levels at the 1 percent level of significance.

Table 4—Insurance Uptake and Enforcement of Eligibility

	Individual (P1)		Household (P3)		Group (P4)	
	Dependents	HH	Dependents	HH	Dependents	HH
PKR 100	0.166	0.410	0.182	0.258	0.167	0.265
	(0.025)	(0.048)	(0.031)	(0.040)	(0.034)	(0.043)
	[577]	[166]	[660]	[186]	[648]	[189]
PKR 90	0.303	0.651	0.420	0.472	0.269	0.332
	(0.026)	(0.037)	(0.042)	(0.040)	(0.039)	(0.041)
	[792]	[235]	[752]	[214]	[897]	[247]
PKR 80	0.341	0.740	0.484	0.510	0.427	0.477
	(0.026)	(0.032)	(0.053)	(0.048)	(0.046)	(0.044)
	[785]	[227]	[741]	[208]	[850]	[239]
PKR 70	0.385	0.776	0.708	0.739	0.656	0.683
	(0.033)	(0.031)	(0.048)	(0.040)	(0.055)	(0.050)
	[827]	[228]	[784]	[222]	[761]	[202]
Observations	2,981	856	2,937	830	3,156	877
F-stat (P1 versus P*i*)			3.84	65.58	8.68	79.84

Notes: Standard errors in parentheses are clustered at the level of the village. Number of dependents (households) are in square brackets. *F*-stat is the *F*-statistics of a joint hypothesis test for equality of take-up rates between policy P1 and policies P3 and P4, respectively.

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.

Online Appendix Table A2 sheds further light on the determinants for households to enroll in the different insurance products. In the individual product (P1), household size seems to play a more limited 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 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 that 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.

V. 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 whether 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.

A. *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 inpatient costs caused by 334 hospital events reported in our phone survey after insurance was introduced. Expected costs are further adjusted to claim behavior because not all hospitalization events lead to a claim. 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 (including health history measured at baseline, as proxied by costs of inpatient and outpatient care, chronic diseases, etc.).[18] We follow this approach for several reasons.

The primary reason is that despite the scale of our RCT, we have too few claim observations, which is a limitation of this study. 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. Studies that only rely on baseline risk measures face uncertainty about how well these 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, as well as an adjustment for reimbursement costs eventually incurred by the insurer. The model used for this mapping explains a significant amount of the variation in expected costs, as indicated by the test for model significance.[19]

In addition, 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.[20] Specifying the cost index as a function of baseline values avoids any such confounding.[21] 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 individuals with the same baseline variables—irrespective of their insurance status. The comparison between insured and noninsured hence remains unbiased. Note that our experimental setup nevertheless 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

[18] See online Appendix D for further details on the parametric models we use to predict the cost index. Note that results are robust to other prediction models.

[19] Refer to online Appendix D for the estimation results and model specification test. Also, note that we do not expect health shocks to be perfectly predictable.

[20] 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.

[21] 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). Online Appendix Table C1 reveals no statistically significant difference in the household size reported at baseline.

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. Online 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.[22] The main analysis uses the predictive model that includes data from all villages in the experiment in order to maximize precision of the estimation.

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 72.29. Online Appendix D documents that the index is balanced between policies P1, P3, and P4. Online Appendix Table D2 furthermore reveals that the insurance scheme pays out less than the claimable amounts suggested by the phone survey data (only about 39 percent on average).[23] Note, however, that the premiums offered are still actuarially fair on average given the real payouts made.

B. *Presence of Adverse Selection: Positive Correlation Test*

As described in Section I, 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 from a bivariate regression of the expected cost index on the binary insurance status for each of the offered policies (and the distribution of the estimated difference between insured and noninsured). 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 percent larger for insured individuals, and the difference is highly significant (p-value $= 0.0002$). 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 percent higher for insured. This difference is statistically significant at the 10 percent level for policy P3 and insignificant for P4.[24]

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

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

[23] We do not know whether this is caused by exaggerated amounts in the phone survey or incomplete utilization of the policy. A lack of awareness and underutilization, however, are not unusual in low-income insurance markets. We leave the discussion about the impact of possible payout problems to a different literature (e.g., Cole et al. 2013; Biener, Landmann, and Santana 2019), as there is no immediate relation to adverse selection.

[24] Note that all of our inference is based on a bootstrap procedure of the whole data-generating process, clustering standard errors at the highest level of randomization (villages).

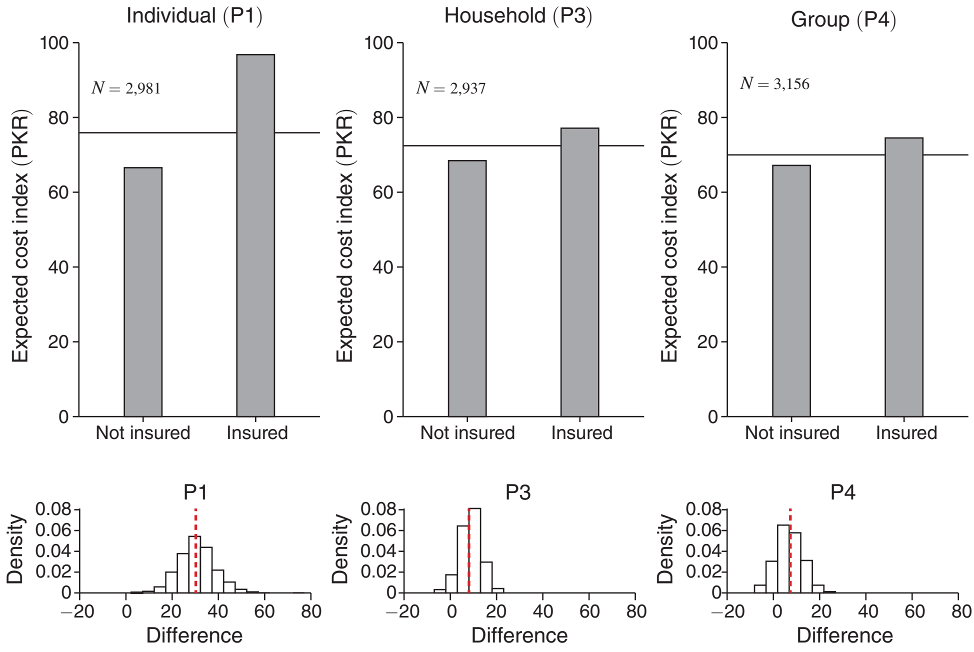

FIGURE 3. POSITIVE CORRELATION TEST: EXPECTED COST INDEX AND TAKE-UP, BY POLICY

Notes: Bars indicate mean values of the health cost index by insurance status and policy. The bootstrap distribution of the mean difference between Not insured and Insured derived from OLS regression of the health risk index on a binary insurance status indicator is shown in the lower panel and accounts for clustering at the village level. The respective bootstrap p-values of the mean difference for P1, P3, and P4 are 0.0002, 0.0534, and 0.1986, respectively.

selection. Note that this pattern can explain the partial insurance uptake within the household established in Section IV. The corresponding demand analysis in online Appendix Table A2 confirms that idiosyncratic health risk factors are a much better predictor for insurance uptake in the individual product 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.[25]

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

[25] We simulate a simple selection process in online Appendix E to disentangle the design effect from the tendency to exploit the scope for selection. We explain more in the next subsection, as the simulation predicts cost curves.

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 online Appendix Table A3, 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 nonhealth variables ("observables") 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 ("unobservables").[26] 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 that is caused by a transition of lower-risk individuals out of the insurance pool as prices increase. The setup of our RCT allows the investigation of such dynamics caused by price changes. We discuss the corresponding analyses in the next section.

C. *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 I, 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. The upper and lower adjacent lines depict the 2.5 and 97.5 percent quantiles of the bootstrap distribution of the expected cost index for a given policy and discount level. For the individual-level policy (P1), the mean costs associated with the insurance pool decrease with demand (i.e., with lower premiums). For the household (P3) and group (P4) policies, there also seems to be a downward shift in the cost distribution with decreasing premiums, but the slope is smaller than under the individual policy (P1).

Table 5, panel A shows the result of testing for a trend in the mean cost index of insured individuals by policy by estimating a linear regression of the individual-specific cost index on aggregate demand for the corresponding policy at the respective price. Findings lack precision, especially when there are fewer observations in the insurance pool, but the downward slope of the average cost curve

[26] The analysis in online Appendix Table A4 confirms that unobservables are very important to understand selection in the individual policy (P1) and furthermore indicates that trying to predict expected costs with observables only might lead to the (flawed) conclusion that selection into insurance is limited and of similar magnitude across policies.

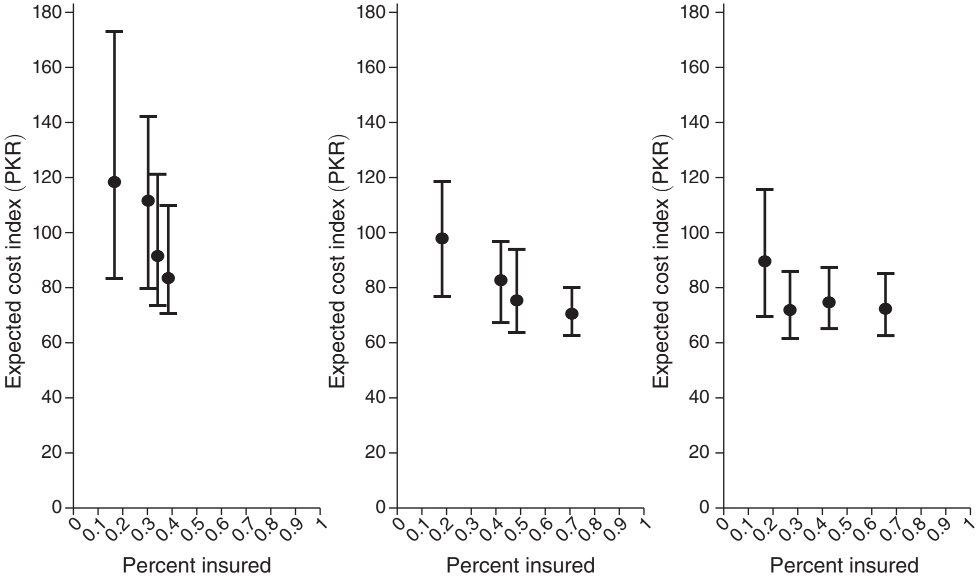

Figure 4. Distribution of Expected Cost Index of Insured over Demand, by Policy

Notes: This figure illustrates shifts in the expected cost distribution by discount level and policy regime. The upper and lower adjacent lines depict the 2.5 and 97.5 percent quantiles of the bootstrap distribution of the expected cost index for a given policy and discount level. Note that this figure abstracts from the variation in uptake induced by bootstrap resampling by depicting demand levels observed in the original sample.

Table 5—Trend in Expected Costs

	P1	P3	P4	P1	P3	P4
	Panel A. Insured			*Panel B. Insured, restricted*[a]		
Uptake (%)	−183.205	−45.950	−15.197	−32.423	−14.064	−9.152
	(112.379)	(19.383)	(17.148)	(8.451)	(5.494)	(6.382)
Constant	156.8931	101.686	81.607	108.333	86.506	79.139
	(38.389)	(11.607)	(8.628)	(9.496)	(5.974)	(6.263)
Observations	922	1,350	1,211	922	1,350	1,211
	Panel C. Noninsured			*Panel D. Noninsured, restricted*[b]		
Uptake (%)	−52.796	16.439	6.107	−32.503	−5.617	−5.767
	(41.939)	(17.075)	(24.661)	(8.328)	(6.032)	(7.398)
Constant	82.446	61.941	65.137	75.910	72.441	69.987
	(13.375)	(7.238)	(8.460)	(3.848)	(3.759)	(3.522)
Observations	2,059	1,587	1,945	2,059	1,587	1945

Notes: Point estimates result from OLS regression of expected cost index on average demand for relevant policy at respective discount. Standard errors (in parentheses) are based on bootstrap resampling and account for clustering at the village level.

[a] The estimation is restricted to pass through the average cost index for the respective policy at a demand level of 0.

[b] The estimation is restricted to pass through the average cost index for the respective policy at full demand.

tends to be stronger in the individual policy (P1) than in the household and group policies (P3, P4). To improve precision, we leverage one necessary attribute of the curve: with 100 percent 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, in line with the analyses in Einav, Finkelstein, and Cullen (2010). Table 5, panel B shows the results. The slope of the cost curve is relatively large and highly significant for the individual policy P1 (−32.423). Figures are smaller for household policy P3 (−14.064, significant at 1 percent level) and group policy P4 (−9.152, insignificant). When comparing the slopes, we find significant differences between P1 and P4 (p-value: 0.0561).

Panels C and D in Table 5 test the relationship between the cost index and the share insured for the noninsured.[27] Results again lack precision in the unrestricted case, but the restricted regression, which exploits that average costs among the noninsured should equal overall average costs when uptake is zero, shows a highly significant negative slope for P1. For this reason, it seems that marginal individuals switching the insurance status to being insured in response to lower prices are high risk relative to those remaining noninsured. This is in line with the economic theory on adverse selection discussed in Section I. In contrast, we do not observe such a pattern for the noninsured in household (P3) and group (P4) policies. Again, this pattern is consistent with adverse selection to exist mainly in the individual policy (P1).

We conduct several robustness checks. For instance, we use an alternative health risk measure that 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.[28] Finally, we validate our analysis by comparing real hospitalization costs, claim incidences and claimed amounts among the insured during the product cycle between policy types. All three measures are significantly higher in the individual policies' insurance pool (see online Appendix Table D2).

Further, we assess the extent to which the lower cost curve slopes in the pooled policies are mechanical by simply reducing the scope for selection. For example, there could be a lower tendency to exploit the scope for selection, because it is more difficult to estimate (relative) expected costs at the household level than for certain individuals within the household. To investigate this question, we simulate perfect and "fuzzy" selection under the different policies. Specifically, we vary the tendency to deviate from perfect cost *ranking*. Our results suggest that when moving from the individual to the household and group policy, cost curves become more and more flat by design. In addition, though, we find more and more deviation from perfect cost *ranking*, suggesting that there is a separate effect from pooling than just to reduce the scope for perfect selection. We provide more details in online Appendix E.

VI. Market Equilibrium and 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

[27] Online Appendix Figure A1 illustrates the corresponding cost points similar to Figure 4.
[28] The results for these robustness checks are available upon request.

have to enroll. This section investigates the market equilibrium and welfare consequences of adverse selection under the different policies. As discussed in Section I, the exogenous price variation induced by the RCT setting identifies both the demand and the average cost curves. To analyze potential market outcomes, we need to connect the demand estimates from Section IV with the analyses on the slope of the average cost curve in Section VC. We use priors to constrain our estimates to exhibit reasonable features. First, as discussed in Section VC, we restrict the average cost curve to pass through the mean of the cost index at full demand. In the linear case, the marginal cost curve can easily be derived afterward ($MC' = 2 \times AC'$). In addition, we restrict the slope of the demand curve to yield full coverage at price zero or above.[29] 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 result of the exercise is shown in Figure 5. It plots the average demand at different premium prices and 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). 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 significantly less steep and more often situated below the demand curve. This leads to higher equilibrium demand 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 online Appendix Table A5), but shifts in the average cost curves (parameter estimates in Table 5) also play a role.

Drawing conclusions about the welfare implications of adverse selection is more complicated. In principle, we could compare how close the policies are to the efficient allocation and compute welfare losses. Neoclassical welfare analysis assumes that willingness to pay (estimated via demand curves) measures utility derived by coverage. In our context, this interpretation might be flawed for several reasons: misconceptions about insurance benefits, liquidity constraints, or simply irrational behavior could affect observed levels of insurance demand.[30] We indeed find uptake patterns consistent with liquidity constraints for household and group policies (see the discussion on demand in Section IV). What is more problematic for a comparative welfare analysis across policies it that demand changes when moving from individual to bundled policies. Our finding of higher *average* willingness to pay

[29] In other words, we assume full take-up if the product was offered for free. While other studies do not always find this to hold (Banerjee et al. 2019), our restriction is binding in only one case (P1) and even there fits the data very well. In addition, we provide market equilibrium and welfare results for the unrestricted case of P1 in Table 6.

[30] The interpretation of the demand curve might also 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.

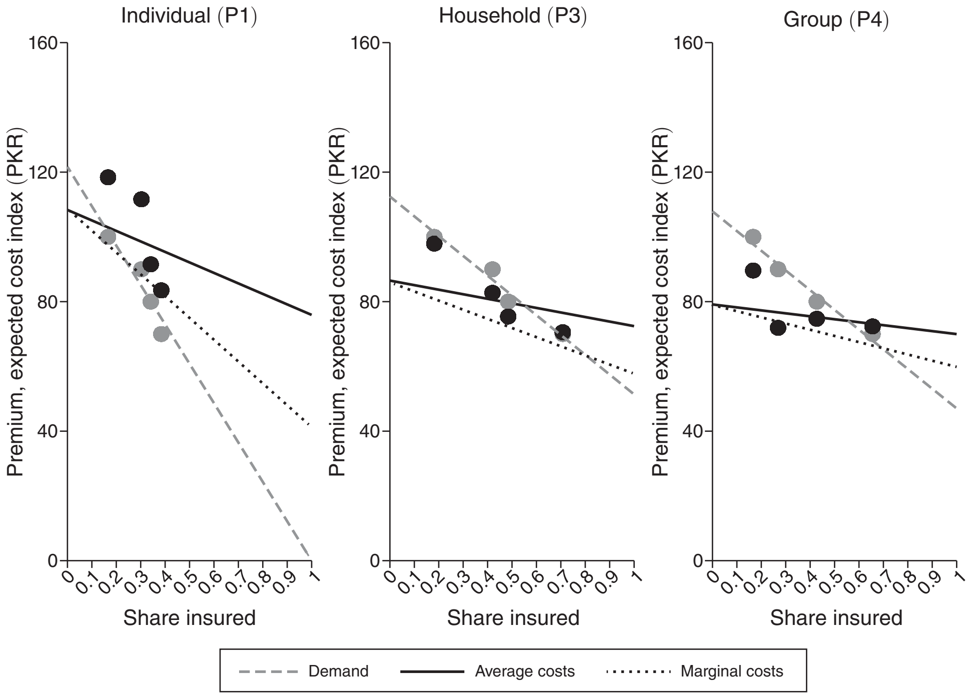

Figure 5. Market Equilibrium and Efficient Allocation, by Policy

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 gray. 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 both curves are shown in Table 5 and online Appendix Table A5 and account for clustering of standard errors at the village level.

for bundled policies is not easy to reconcile with simple neoclassical theory under perfect information. In such an environment, the demand curve might become more flat when moving from individual to bundled policies, but average willingness to pay should remain similar.[31] We provide welfare estimates in the following, but it is important to keep these reservations in mind, in particular when comparing the individual policy P1 with bundled policies P3 and P4.

Table 6 shows the equilibrium and the efficient allocations under the different policies and calculates the resulting welfare. Welfare estimates are much higher under bundled policies, both in equilibrium and in the efficient allocation. The latter is especially inconsistent with the economic intuition that welfare should be highest under efficient allocation of the individual policy. This counterintuitive result follows

[31] Formal proofs for this intuition involve assumptions, of course. For example, assuming constant absolute risk aversion and independent health shocks within a unitary household model, it is possible to show that the sum of the willingness to pay for each individual household member as indicated by our demand curve is equal to the willingness to pay for the whole household.

Table 6—Market Equilibrium and Welfare Analysis

	Individual (P1)	Individual (P1)[a]	Household (P3)	Group (P4)
Equilibrium				
Price	103.41	101.78	78.87	73.81
Quantity	0.15	0.19	0.55	0.56
Welfare	1.36	2.65	9.29	9.43
Efficient				
Price	93.67	91.20	64.2	66.50
Quantity	0.23	0.26	0.79	0.68
Welfare	1.56	2.84	10.22	9.71
Loss				
Quantity	0.08	0.07	0.24	0.12
Welfare	0.20	0.19	0.93	0.28
% welfare	0.84	0.71	1.45	0.55

Notes: % relative to cost of coverage in the efficient allocation (see Bundorf, Levin, and Mahoney 2012).
[a] Unconstrained results (allowing less than full demand at zero price).

from the higher average willingness to pay under bundled policies. To assess the effects of adverse selection, we look at the difference between equilibrium and efficient allocation. Losses in quantity caused by adverse selection are higher for bundled policies (0.12–0.25) than for the individual policy (0.08). Further, the calculated welfare loss is higher for the household and group insurance (P3: 0.93, P4: 0.28) than in the individual insurance policy (P1: 0.20). 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. Following Bundorf, Levin, and Mahoney (2012), we also calculate the welfare loss relative to cost of efficient coverage (average cost × quantity). We find relative welfare losses of 0.84 percent for P1, 1.45 percent for P3, and 0.55 percent for P4. Even though the comparability between individual and bundled policies is limited for the reasons described above, these results suggest that bundling has mixed effects regarding the relative welfare loss.

The welfare results should be interpreted with additional 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. Online Appendix Figure A2 suggests a more marked contrast between individual (P1) and bundled policies (P3, P4): the market for individual policies breaks down completely.[32] We therefore interpret the linear specification as a conservative estimate of the difference among policies.

[32] 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.

VII. Discussion and Conclusion

This paper provides robust evidence on adverse selection in low-income health insurance markets. We analyze an RCT 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, estimate how selection changes at different points of the price curve, and test different mechanisms against adverse selection. Our analysis of adverse selection is based on individual health characteristics at baseline that 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 and 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 microfinance 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. When calculating relative welfare costs of adverse selection within each policy (following Bundorf, Levin, and Mahoney 2012), results are more mixed, though.

The conclusions related to welfare are subject to some reservations. 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 due to a range of possible frictions (Spinnewijn 2017; Handel, Kolstad, and Spinnewijn 2019). 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 focuses on simple pooling products. This means that only one policy is offered and no additional measures against adverse selection, such as copayments or ex ante screening, are included. Our results show that even under these circumstances, household policies might achieve a sustainable pool of insurance clients with willingness to pay exceeding costs. This is an interesting aspect, as in other contexts (such as insurance for low-income US adults in Finkelstein, Hendren, and Shepard 2019, or Indonesian households in Banerjee et al. 2019), it is difficult to generate demand even with high subsidies. High demand in our case might be explained by the lack of alternative protection, the nonexistence of lower-cost alternatives for inpatient services, the severe consequences of financial shocks, and high trust in the local NGO distributing the insurance. In any case, our findings are good news for organizations interested in patching missing 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." *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): 941–73.

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 and Proceedings* 104 (5): 291–97.

Banerjee, Abhijit, Amy Finkelstein, Rema Hanna, Benjamin A. Olken, Arianna Ornaghi, and Sudarno Sumarto. 2019. "The Challenges of Universal Health Insurance in Developing Countries: Evidence from a Large-Scale Randomized Experiment in Indonesia." NBER Working Paper No. 26204.

Biener, Christian, Andreas Landmann, and Maria Isabel Santana. 2019. "Contract Nonperformance Risk and Uncertainty in Insurance Markets." *Journal of Public Economics* 175: 65–83.

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): 411–40.

Bundorf, M. Kate, Jonathan Levin, and Neale Mahoney. 2012. "Pricing and Welfare in Health Plan Choice." *American Economic Review* 102 (7): 3214–48.

Cabral, Marika. 2017. "Claim Timing and Ex Post Adverse Selection." *Review of Economic Studies* 84 (1): 1–44.

Cameron, A. C., P. K. Trivedi, Frank Milne, and J. Piggott. 1988. "A Microeconometric Model of the Demand for Health Care and Health Insurance in Australia." *Review of Economic Studies* 55 (1): 85–106.

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): 101–06.

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

Cole, Shawn, Xavier Giné, Jeremy Tobacman, Petia Topalova, Robert M. Townsend, and James Vickery. 2013. "Barriers to Household Risk Management: Evidence from India." *American Economic Journal: Applied Economics* 5 (1): 104–35.

de Meza, David, and David C. Webb. 2017. "False Diagnoses: Pitfalls of Testing for Asymmetric Information in Insurance Markets." *Economic Journal* 127 (606): 2358–77.

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

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.

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. 2019. "Subsidizing Health Insurance for Low-Income Adults: Evidence from Massachusetts." *American Economic Review* 109 (4): 1530–67.

Fischer, Torben, Markus Frölich, and Andreas Landmann. 2023. "Replication data for: Adverse Selection in Low-Income Health Insurance Markets: Evidence from an RCT in Pakistan." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.38886/E178421V1.

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.

Hackmann, Martin B., Jonathan T. Kolstad, and Amanda E. Kowalski. 2015. "Adverse Selection and an Individual Mandate: When Theory Meets Practice." *American Economic Review* 105 (3): 1030–66.

Handel, Benjamin R., Jonathan T. Kolstad, and Johannes Spinnewijn. 2019. "Information Frictions and Adverse Selection: Policy Interventions in Health Insurance Markets." *Review of Economics and Statistics* 101 (2): 326–40.

Heltberg, Rasmus, and Niels Lund. 2009. "Shocks, Coping, and Outcomes for Pakistan's Poor." *Journal of Development Studies* 45 (6): 889–910.

ILO Microinsurance Innovation Facility. 2014. *ILO's Microinsurance Innovation Facility Annual Report 2013*. Geneva, Switzerland: International Labour Organization.

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.

Lammers, Judith, and Susan Warmerdam. 2010. "Adverse Selection in Voluntary Micro Health Insurance in Nigeria." Unpublished.

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.

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 and Medicine* 71 (12): 2074–82.

NRSP. 2016. "Client-Level Information from the Management Information System." National Rural Support Programme of Pakistan.

Pakistan Ministry of Health. 2009. *National Health Policy 2009: Stepping towards Better Health* (*Draft 19 Feb 2009*). Islamabad: Pakistan Ministry of Health.

Panhans, Matthew. 2019. "Adverse Selection in ACA Exchange Markets: Evidence from Colorado." *American Economic Journal: Applied Economics* 11 (2): 1–36.

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

Rural Support Programmes Network. 2015. *Outreach #26*. Islamabad: Rural Support Programmes Network.

Spinnewijn, Johannes. 2017. "Heterogeneity, Demand for Insurance, and Adverse Selection." *American Economic Journal: Economic Policy* 9 (1): 308–43.

Swiss Re. 2010. *Microinsurance—Risk Protection for 4 Billion People*. Zürich, Switzerland: Swiss Reinsurance Company Ltd.

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: World Bank.

World Bank. 2012. *Micro Insurance in Pakistan: A Diagnostic Study*. Washington, DC: World Bank.

Yao, Yi, Joan T. Schmit, and Justin R. Sydnor. 2017. "The Role of Pregnancy in Micro Health Insurance: Evidence of Adverse Selection from Pakistan." *Journal of Risk and Insurance* 84 (4): 1073–1102.

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