Moral hazard and selection among the poor: evidence from a randomized experiment

Spenkch, Jörg L.

University of Chicago

November 2011

Moral Hazard and Selection Among the Poor: Evidence from a Randomized Experiment*

Jörg L. Spenkuch

University of Chicago

First Draft: January 2011

This Version: November 2011

Abstract

Not only does economic theory predict high-risk individuals to be more likely to purchase insurance, but insurance coverage is also thought to crowd out precautionary activities. In spite of stark theoretical predictions, there is conflicting empirical evidence on adverse selection, and evidence on ex ante moral hazard is very scarce. Using data from the Seguro Popular Experiment in Mexico, this paper documents patterns of selection on observables into health insurance as well as the existence of non-negligible ex ante moral hazard. More specifically, the findings indicate that (i) agents in poor self-assessed health prior to the intervention have, all else equal, a higher propensity to take up insurance; and (ii) insurance coverage reduces the demand for self-protection in the form of preventive care. Curiously, however, individuals do not sort based on objective measures of their health.

* I would like to thank Amy Finkelstein, Roland Fryer, Steven Levitt, Bruce Meyer, and David Toniatti as well as seminar participants at TU Dortmund for helpful comments and discussions. I am also grateful to Gary King and coauthors for making their data from the Seguro Popular Assessment available. Magdalena Balibrea and Steven Castongia provided excellent research assistance. A portion of this paper was written while visiting the Rheinisch-Westfälisches Institut für Wirtschaftsforschung in Essen, Germany. Financial support from the German National Academic Foundation is gratefully acknowledged. All views expressed in this paper as well as any remaining errors are solely my responsibility.
Correspondence can be addressed to the author at Department of Economics, University of Chicago, 1126 E 59th Street, Chicago, IL 60637, or by e-mail: jspenkuch@uchicago.edu.

I. Introduction

Starting with the seminal work of Arrow (1963), economists have long been aware of the possibility of market failure due to asymmetric information. Standard explanations for why the First Welfare Theorem may not hold rely on adverse selection (Akerlof 1970, Rothschild and Stiglitz 1976) as well as ex ante and ex post moral hazard (Ehrlich and Becker 1972, Pauly 1968, 1974).

In the context of insurance markets, these models postulate a positive correlation between coverage and risk occurrence (Chiappori and Salanié 2000, Chiappori et al. 2006).[1] Yet, empirical tests of this prediction point often in conflicting directions. Cawley and Philipson (1999), for instance, find little or no information asymmetries in the market for life insurance, whereas Finkelstein and Poterba (2002, 2004) conclude the opposite with respect to that for annuities.[2]

However, even if 'high risk' agents are, ceteris paribus, more likely to purchase insurance than their low risk counterparts, insurance status and ex post risk occurrence need not be correlated when individuals possess multidimensional private information (de Meza and Webb 2001); and a positive correlation alone is uninformative about whether market failures are due to adverse selection or moral hazard. But exactly this distinction is important for public policy. Consider, for example, the case of insurance mandates—as implemented in many European countries. While mandates are often welfare enhancing when inefficiencies are due to selection, it is well known that requiring coverage may actually decrease welfare in the presence of moral hazard.

Using data from a large-scale randomized field experiment in Mexico, this paper makes two contributions to the literature. First, it exploits rich information on agents' health prior to the intervention coupled with experimental variation in access to health insurance in order to analyze sorting based on risk, thereby disentangling moral hazard from selection. Second, the paper explicitly tests for ex ante moral hazard.[3]

In 2004 the Mexican government introduced the System for Social Protection in Health (Sistema de Protección Social en Salud, SPSS). This reform sought to protect workers outside the formal sector from catastrophic health expenditures through the introduction of Seguro Popular en Salud (SP). By the letter of the law SP constitutes a wealth transfer program and was not explicitly designed as an insurance scheme. Economically, however, it serves the function of a voluntary health insurance option available to

[1] In the insurance setting adverse select refers to the idea that 'high risk' individuals are more likely to buy insurance, or purchase more coverage than 'low risk' ones. Ex ante moral hazard is said to exist when the insured invest less effort in self-protection than their uninsured counterparts; and ex post moral hazard captures agents' failure to fully internalize the cost of covered services.

[2] For a useful review of the empirical literature see Chiappori and Salanié (2003).

[3] Unfortunately, the data only contain information on out-of-pocket (as opposed to total) medical expenditures, and do not lend themselves to a test of ex post moral hazard. See Cutler and Zeckhauser (2000) for a survey of earlier, non-experimental attempts to infer adverse selection from the correlation between insurance status and individual characteristics proxying for risk.

the uninsured and their dependents, who in 2004 comprised about half of the Mexican population. That is, to obtain full coverage for a large number of health interventions and pharmaceuticals families must formally affiliate with SP and those in upper four quintiles of the income distribution pay a small, progressive income based fee.

Due to administrative and budgetary constraints SP has been rolled out in different stages over a multi-year horizon. Commissioned by the government to evaluate SP, King et al. (2007, 2009a) use the staged rollout to experimentally introduce geographic variation in the availability of health insurance for the poor. They report that within 10 months after random assignment SP reduced catastrophic out-of-pocket health expenditures by 6.5 percentage points, but did not increase the utilization of medical care (King et al. 2009a). All data used in this paper stem from their intervention. Since the Seguro Popular Experiment induced a large number of individuals to obtain insurance coverage and collected information on the utilization of preventive care as well as on agents' health *prior* to treatment, it is exceptionally well suited for studying selection and ex ante moral hazard.[4]

As predicted by theory, the results show that 'high risk' agents are, *ceteris paribus*, more likely to opt into SP—although the insured are not more 'risky' *on average*. That is, despite the absence of a positive raw correlation between agents' insurance status and proxies of risk, this paper presents evidence of the systematic selection predicted by theory. In particular, individuals who rated their health as "bad or very bad" before SP became available are 6.9 percentage points more likely to sign up for SP than those in "good or very good" health (compared to an overall treatment effect of 29 percentage points). Curiously, however, agents in the experiment sort only on pre-period medical expenditures and subjective well-being. There appears to be no selection on objective measures of health—possibly because individuals are less aware of the latter.

Relying on individuals' utilization of preventive medical care as a proxy for self-protection, this paper also demonstrates the existence of considerable ex ante moral hazard. Although the full-cost of preventive care decline with insurance coverage, the effect of SP on the utilization of these services is *negative* and non-trivial in size. Given the positive price effect, such a decline is likely due to ex ante moral hazard.

The results presented in this paper speak directly to those in the recent literature. Finkelstein and McGarry (2006) present evidence of offsetting multidimensional private information in the market for long-term care insurance; and Fang et al. (2008) show that the insured in the Medigap market spend *less* on health care, which the authors attribute to neutralizing advantageous selection on education and cognitive ability. The present paper identifies risk based sorting from experimental variation, but does not

[4] King et al. (2009a) do not test for selection on observable risk factors and do not relate the use of preventive care to ex ante moral hazard.

find advantageous selection on education. Notwithstanding a positive raw correlation between education and insurance status, the causal effect of education on insurance take up goes in the opposite direction—at least in this very specific setting. Unfortunately, the data do not point to other dimensions of private information that would explain why individuals do not appear to be negatively selected on average.

Although insurance coverage had long been predicted to crowd out precautionary effort (cf. Ehrlich and Becker 1972, Pauly 1974), empirical evidence on ex ante moral hazard has hitherto been very sparse. For the health insurance market Zweifel and Manning (2000) review a number of studies estimating the price elasticity of preventive care. They conclude that "the limited available evidence indicates that the demand for preventive care is a declining function of out-of-pocket money price" (p. 420), but that it is unclear whether or not it is more elastic than that for regular ambulatory care.[5] Evidence on the extensive margin is even scarcer. The findings in Card et el. (2008) suggest no reduction in the demand for preventive services due to Medicare coverage, but Klick and Stratman (2007) argue that health insurance coverage is associated with increases in BMI.[6]

Taken at face value the results of this paper have potentially important policy implications. If ex ante moral hazard is indeed a non-trivial force in the market for health insurance, then the welfare consequences of insurance mandates are theoretically ambiguous—even if agents are strongly adversely selected and if preferences were homogenous. Moreover, optimal insurance contracts would not only trade off risk sharing and ex post moral hazard through co-payments (Zeckhauser 1970), but would also have to provide appropriate incentives to engage in precautionary activities. But as self-protection resembles the classic multi-task principal-agent problem (Holmström and Milgrom 1991), it is a priori unclear what the optimal incentive scheme might look like.

The remainder of the paper proceeds as follows. Section II provides background information on the health care sector in Mexico as well as on the experiment conducted by King et al. (2007, 2009a). Section III presents the empirical results, and the last section concludes.[7]

II. Background Information and Data Description

A. Institutional Background

The Mexican health care system is characterized by several vertically integrated social security institutions (e.g., IMSS, IMSSTE, PEMEX, SEDENA, etc.).[8] They provide health insurance and medical

---

[5] For estimates from the RAND Health Insurance Experiment see Lillard et al. (1986) and Newhouse et al. (1993).
[6] In other markets the evidence on ex ante moral hazard has been ambivalent as well. See, for instance, see Abbring et al. (2003) or Abbring et al. (2008).
[7] There are three appendices. Appendix A contains a formal proof omitted from the body of the paper, and Appendix B discusses selection on education and pre-period medical expenditures. A Data Appendix with the precise definitions and sources of all variables used in the analysis is provided on the author's website.
[8] The following summary draws heavily on the description in OECD (2005).

care to workers in the formal sector of the economy as well as to their dependents. Generally, facilities run by a social security provider offer free treatment to any of its affiliates; and membership in these institutions is mandatory for all salaried workers. However, given the size of the informal sector, only about 50% of the Mexican population is covered by the social security system. The other half remains without health insurance (OECD 2005).[9]

The uninsured purchase medical care either in a large, unregulated private market or in (often inadequately funded) public clinics and hospitals. Although clinics operated by the Ministry of Health or State Health Services offer medical care at below full-cost prices, private out-of-pocket payments account for c. 55% of all health related spending in Mexico (OECD 2005). Consequently, catastrophic health expenditures and poverty go often hand in hand.

In January 2004 the Mexican government introduced the System for Social Protection in Health (Sistema de Protección Social en Salud, SPSS). Over a multi-year horizon this reform aims at eliminating catastrophic health expenditures for workers outside the formal sector, and at improving public health care provision by injecting new resources into the system as well as re-balancing existing transfers from the federal government to the states. Ultimately, SPSS is projected to increase public spending on health care by .8–1.0 percentage points of GDP—roughly 30% of the 2002 level (OECD 2005).[10]

At the heart of the reform is Seguro Popular en Salud (SP). It is important to recognize that *formally* SP was intended to be wealth transfer program, not an insurance scheme. That is, premiums are not risk adjusted, and are not set to break even. Nevertheless, *economically* SP resembles a voluntary health insurance option available to all uninsured and their dependents, irrespective of pre-existing conditions. In order to obtain coverage for over 250 health interventions and more than 300 pharmaceuticals (without any co-pay) families must formally affiliate with SP and pay a small, progressive income based fee.[11] Poor households, i.e. those in the lowest quintile of the income distribution, are exempt from paying premiums. The same is true for participants in the Oportunidades anti-poverty program, who are automatically enrolled in SP by the states.[12]

By the letter of the law every citizen who is not covered by the social security system is allowed to enroll in SP. However, financial and administrative constraints require the program to be rolled out in multiple stages, and only to areas whose health facilities meet a set of minimal requirements. At least in theory, individuals living in regions in which SP has not yet been introduced can travel to the nearest receiving area in order to affiliate. Yet, the benefits of doing so may be rather small in practice, as sick

---

[9] Only 3% of individuals buy private insurance, many of whom are already affiliated with a social security provider (OECD 2005).

[10] See, for instance, Frenk et al. (2006) for details on SPSS; or Homedes and Ugalde (2009) for a history of health reform in Mexico.

[11] The services covered by SP treat illnesses accountable for 95% of the disease burden in Mexico (King et al. 2007).

[12] In the data 93.6% of SP affiliates report paying no premium.

individuals would have to go back to take advantage of the free treatment to which their SP affiliation entitles them. It is, therefore, useful to think of the staged introduction of SP as introducing geographic variability in poor families' access to affordable health insurance.

Several observational studies evaluate SP's success in reducing out-of-pocket expenditures and improving access to medical care, generally finding positive effects (e.g., Barros 2008, Bleich et al. 2007, Gakidou et al. 2006, Knox 2008, and Sosa-Rubi et al. 2009). However, King et al. (2007) report that political considerations played an important role in determining which regions received SP first, thereby raising concerns about identifications strategies based on non-experimental variation.

B. Data Description

To assess the effectiveness of SP the Mexican Ministry of Health commissioned a large-scale randomized field experiment implemented and analyzed by King et al. (2007, 2009a, 2009b). They exploit the staged rollout to *experimentally* introduce geographic variation in the availability of SP. All data used in this paper stem from their program evaluation.[13]

In a first step King and coauthors partitioned Mexico into 12,284 "health clusters", defined as the catchment area around an existing or planned health clinic. As 19 of 32 states refused to take part in the evaluation, a subset of 7,078 (5,439 rural and 1,639 urban) clusters remained ultimately eligible for possible participation in the experiment.[14] To assure that the treatment and control group are balanced on observables, eligible health clusters were matched into pairs. Based on closeness of match and likelihood of compliance a subset of 74 pairs was selected to participate; and assignment to the treatment or control group was based on the flip of a fair coin *within* each pair. Figure 1 shows the location of treatment and control clusters.

Concurrent with random assignment (August – September 2005) two baseline surveys collected data on households and health facilities. Besides standard demographic variables and information on the utilization of medical care, the household questionnaire also elicited out-of-pocket expenditures and self-assessed health prior to the intervention. In addition, the data also contain objective measures of individuals' health from anthropomorphic tests administered on site, e.g., BMI, cholesterol level, or systolic blood pressure.

Unfortunately, budget constraints limited the number of clusters in which the household survey could be conducted to 50 pairs. Consequently, the analysis in this paper only uses data on the 32,426 respondents to the household survey with non-missing information on insurance status.

---

[13] The following description of the experiment borrows heavily from King et al. (2007, 2009a). For additional details the interested readers should consult the aforementioned studies.

[14] Clusters in which a substantial fraction of families had already been affiliated with SP had been discarded.

Table 1 displays summary statistics from the baseline survey.[15] Obviously, individuals in the Seguro Popular Experiment are not a random sample of Mexicans. Given that the experiment intentionally focused on rural areas with low insurance coverage, it is not surprising that only 10.6% of participants are urban dwellers, and a mere 21.8% have any health insurance at baseline. Comparing individuals in the Seguro Popular Experiment to the nationally representative sample of the ENSANut 2005 (Encuesta Nacional de Salud y Nutrición), King et al. (2007) report that the demographics of both populations look otherwise very similar.

More importantly, on almost all observable characteristics the treatment and control group appear roughly balanced. Yet, there are two important exceptions. Individuals in the treatment group are more likely than those in the control group to have had a mammogram in the pre-period ($p = .04$) and are more frequently already covered by insurance ($p = .05$). Although true random assignment does not seem implausible (with 2 out of the 33 displayed pairwise comparisons being statistically at the 5% level), the summary statistics in Table 1 suggest that it might nevertheless be important to account for chance differences at baseline. To this end Section IV.C performs a series of robustness checks finding little indication that the results reported in this paper are driven by such differences.

Actual treatment consisted in a local media campaign to inform families about the benefits of SP and to encourage them to enroll, as well as in setting up a 'service and orientation stand' (Módulo de Atención y Orientación) in each treatment cluster so citizens could formally affiliate with SP. Moreover, states attempted to improve health facilities, the availability of drugs, and to hire additional medical personnel.

King et al. (2009a), however, argue that, while the no-charge policy became effective immediately after affiliation, the assessment period of 10 months was likely too short for states' efforts to enhance the existing infrastructure to take effect. For instance, increasing the supply of pharmaceuticals required an open bidding process that took 6–8 months to complete. Similar arguments surely apply to building and accrediting new medical facilities, and to training additional personnel. Therefore, it seems unlikely that access to health care and the quality thereof improved significantly relative to control group.[16]

In order to assess the effect of SP on households' medical expenditures and utilization of services, a follow-up survey successfully re-interviewed 29,897 of the original households during July and August 2006, i.e. approximately 10 months after random assignment.

Table 2 displays individuals' insurance status by experimental assignment and period. The last column shows difference-in-differences estimates (and their standard errors) for each health insurance option. As intended, treatment significantly increased the share of 'insured' individuals relative to the

[15] Information on household characteristics and medical expenditures was elicited from the 'household informant', whereas all individual level information pertains to a household member randomly selected with a Kish table.

[16] After the intervention 69.1% of respondents in the treatment group report being satisfied with the services of their health care provider. The corresponding value for the control group is 68.9%.

control group, in particular the share of SP affiliates. The estimates in Table 2 also suggest a lack of substitution effects—possibly due to the short duration after treatment. That is, there is no evidence that a significant number of previously insured individuals left their provider and affiliated with SP instead.

The fact that the Seguro Popular Assessment induced experimental variation in health insurance coverage among the poor, coupled with rich pre-period information makes it ideally suited for studying selection and ex ante moral hazard.

III. Selection and Ex Ante Moral Hazard: Experimental Estimates

A. Selection on Observables

In particular, the Seguro Popular Experiment admits an unusually clean test for systematic sorting. Since SP draws almost exclusively from the pool of uninsured individuals and the data contain health status and out-of-pocket medical expenditures prior to the intervention, it is not necessary to rely on the commonly used 'positive correlation test' (Chiappori and Salanié 2000). Instead of inferring (the absence of) selection from the contemporaneous correlation between insurance status and realized expenditures (which, as is well known, does not rule out moral hazard), the Seguro Popular Experiment allows to test for selection on riskiness by relating insurance take up in the post-period to a proxy for an individual's inherent risk collected in the pre-period. Intuitively, if selection were important, one would expect a disproportionate number of 'high risk' agents to affiliate with SP as it is being rolled out. While it is, of course, important to know whether the insured are *on average* more risky than the uninsured, a clean test of the theory requires isolating the *causal* effect of an individual's risk type. Moreover, knowledge of the causal effect (as opposed to the mere presence) of selection allows an assessment of whether it is quantitatively important.

To this end, consider the following empirical framework:

(1) $$h_{i,T} = \alpha + \rho X_{i,T-1} + \beta TREATMENT_{i,T} + \gamma TREATMENT_{i,T} X_{i,T-1} + \boldsymbol{Q}_{i,T-1}{}'\boldsymbol{\xi} + \epsilon_{i,T},$$

where $TREATMENT_{i,T}$ indicates whether individual i has received treatment at time T, $X_{i,T-1}$ is a pre-treatment proxy for i's inherent 'riskiness', and $h_{i,T}$ denotes her health insurance status in the post-period.[17] $\boldsymbol{Q}_{i,T-1}$ is a vector of baseline controls containing: gender, age, educational achievement, total household expenditures, a proxy for asset holdings, the number of doctors and nurses in the area, as well

[17] That is, $h_{i,T}$ is equal to one if i is covered by any health insurance at the time of the follow-up survey and zero otherwise. An alternative specification would have been to set $h_{i,T}$ equal to one only if i is affiliated with SP, thereby grouping uninsured individuals and those with formal sector insurance together. Since the paper is interested in testing the predictions of economic theory in this mostly poor and overwhelmingly rural environment, and given that from an economic point of view SP and mandatory health insurance are much more alike than 'formal sector insurance' and 'no insurance', this paper groups 'insured' and 'uninsured' individuals together. It may be comforting to know, however, that the results are qualitatively robust to this modeling choice, which may not be too surprising given the fact that the intervention introduces hardly any variation in mandatory insurance membership (cf. Table 2).

as indicator variables for whether the household participates in Oportunidades, whether it is located in an urban area, and whether individual i belongs to Mexico's indigenous population. Covariates are included to improve precision and to account for chance differences between treatment and control group.

Note that $\hat{\gamma}_{OLS}$ identifies differences in treatment effects between individuals of different risk types. In the context of the Seguro Popular Experiment, the least squares estimate of γ indicates whether 'high risk' individuals are *more* likely than their 'low risk' counterparts to obtain health insurance coverage as it becomes available (independent of other unobserved characteristics). Hence, one can reject the null hypothesis of no selection on $X_{i,T-1}$ if $\hat{\gamma}_{OLS}$ is positive (or negative) and statistically significant.[18]

Tables 3 and 4 display estimates of specification (1) using different pre-period proxies for agents' inherent 'riskiness': self-assessed well-being, and a set of objective measures of health. The rational underlying the choice of these proxies is that agents in poor health before the introduction of SP are also more likely to incur adverse health shocks afterwards, and can therefore be considered 'more risky'.

The right column of Table 3 shows that individuals who assessed their own health as 'good or very good' on the baseline survey are 5.5 and 6.9 percentage points less likely to be induced to take up health insurance than their counterparts in 'fair' and 'bad or very bad' health, respectively. These estimates are large in a real world sense, and they are statistically significant at conventional levels.

Yet, as shown in the column on the left, the raw correlation between health insurance status in the post-period and pre-period health assessments is negative. That is, individuals who rated themselves as healthier, or who did not incur medical expenditures have a higher propensity to be covered by health insurance after the introduction of SP. In principle this could be due to multiple dimensions of private information resulting in net advantageous selection (as documented for different markets by Finkelstein and McGarry 2006, and Fang et al. 2008).

Contrary to the findings of Fang et al. (2008), however, there is no evidence for advantageous selection on education into SP (see Appendix B). Yet, an equally likely explanation in this context may be reverse causality, i.e. healthy individuals might have been more likely to have had health insurance at baseline (for instance, because they are more likely to work in the formal sector of the economy) and continue to do so in the post-period. In fact, when excluding individuals who were covered by health insurance in the pre-period, the coefficients on $X_{i,T-1}$ in Table 3 become very close to zero and are not statistically significant (not reported here).

---

[18] For a formal proof that γ in equation (1) is well identified in the sense that $plim\ \hat{\gamma}_{OLS} = \gamma$ see Appendix A. Also note that despite the inclusion of $\boldsymbol{Q}_{i,T-1}$, the error term $\epsilon_{i,T}$ is almost certainly not orthogonal to $X_{i,T-1}$. For instance, risk aversion and an individual's 'risk type' might not only be correlated with insurance status, but also with each other. Therefore, $\hat{\rho}_{OLS}$ is not well identified, and depending on the correlation between $\epsilon_{i,T}$ and $X_{i,T-1}$, a naïve test of whether 'high risk' agents are more likely to be insured might produce both Type I and Type II errors. Put differently, since $\hat{\rho}_{OLS}$ is not well identified, *levels* of insurance uptake are not useful in gauging the extent of selection. *Differences* in levels, however, are identified and may be used to draw inferences about systematic sorting.

Curiously, there is also no indication of selection on objective measures of health. Table 4 reports estimates of (1) proxying for agents' riskiness with indicator variables for whether their BMI, cholesterol, blood pressure, or glycated hemoglobin (HbA1c)—a criterion for the diagnosis of diabetes—were within their respective 'normal' ranges when measured at the time of the baseline survey.[19] Not only are the estimates of γ in Table 4 not statistically significant and (with one exception) much smaller in magnitude than their counterparts in Table 3, but three of them also carry the 'wrong' sign.

It therefore appears that agents in the Seguro Popular Experiment sort on subjective health assessments, but not on objective ones. One, admittedly unsatisfying, explanation is that the poor might not be fully aware of having elevated levels of cholesterol or HbA1c, whereas they are mindful of 'how they feel'. In such a world only the latter factor would be used in deciding whether or not to sign up for SP, i.e. obtain health insurance coverage. In line with this hypothesis is the fact that BMI, blood pressure, cholesterol, and HbA1c explain less than 1% of the variation in self-assessed health, and that King et al. (2007) report individuals in the experiment viewed the medical tests as incentive to participate in the survey. However, the evidence in favor of this explanation is only circumstantial.[20]

B. Ex Ante Moral Hazard

Apart from adverse selection the theory of asymmetric information also predicts insurance coverage to crowd out self-protection (cf. Ehrlich and Becker 1972, Pauly 1974). Unfortunately, pure measures of self-protection are rarely available, and the Seguro Popular Experiment is no exception. However, the data do contain information on utilization of preventative medical care, i.e. whether an individual had a flu shot, pelvic exam, pap smear, mammogram, or had her eyes checked within the last year. Despite the fact that the full cost of these services decrease with coverage, the insured might still reduce their demand for prevention if ex ante moral hazard is severe enough. Thus, relating insurance status to the utilization of preventive services provides a very strong one-sided test for the presence of ex ante moral hazard.[21]

However, with different imperfect measures of self-protection at hand it is important to efficiently aggregate the available information in order to avoid the multiple inference problem and to improve statistical power to detect effects that go in the same direction. One appealing and particularly simple approach is to calculate average standardized treatment effects (ASTE). E.g. for some set of parameter estimates $\{\hat{\varphi}_s\}$ define

[19] The ranges considered 'normal' for the purposes of this paper are: 18.5 to 25 kg/m^2 for BMI, less than 120 mmHg for systolic blood pressure, less than 200 mg/dL for cholesterol, and less than 6% for HbA1c (see WHO 2000, and NIH 1998, 2002, 2004).

[20] Empirically, self-assessed health is a somewhat better predictor of future health care expenditures than BMI, blood pressure, cholesterol, and HbA1c together. However, the explained share of variation is in both cases very low.

[21] It is important to point out that King et al. (2009a) do analyze the utilization of preventive care, but do not make the connection between utilization of preventive services and ex ante moral hazard and do not account for the problem of multiple inference.

(2) $$ASTE \equiv \frac{1}{|S|}\sum_{s\in S}\frac{\hat{\varphi}_s}{\hat{\sigma}_s},$$

where S denotes the set of available measures, and $\hat{\sigma}_s$ is the standard deviation of outcome $y_{s,i,T}$ in the control group.

To see why calculating ASTEs alleviates the problem of multiple inference note that the ASTE provides a unidimensional measure of the effect of treatment on self-protection. By aggregating information in this way there remains only one hypothesis to be tested, i.e. whether $ASTE$ is statistically different from zero. Moreover, aggregating information (within similar domains) improves efficiency by smoothing out noise.[22]

In calculating standard errors one needs to account for the covariance structure of the individual point estimates, i.e. $\{\hat{\varphi}_s\}$. In order to also cluster on the level of randomization, thus correcting for the correlation of outcomes of individuals within the same health clusters, this paper uses the block bootstrap with 1,000 iterations. That is, standard errors are obtained by repeatedly resampling health clusters from the original data (with replacement), estimating $ASTE$, and calculating the standard deviation of the resulting estimates (see Horowitz 2001 for additional details on the bootstrap).

Notice, to make effect sizes comparable $ASTE$ standardizes each coefficient. Also, in contrast to other means of aggregating information which produce an index-like measure for *each* individual, i.e. factor or principal component analysis, calculating standardized treatment effects does not require valid information on all outcomes for a given observation (or imputation thereof). In the present context this is particularly attractive, since three measures of self-protection are only relevant to females, i.e. pelvic exams, pap smears, and mammograms.

Of course, testing for ex ante moral hazard requires estimates of the causal effect of insurance coverage. Since coverage is almost certainly endogenous, ordinary least squares estimates are likely biased. However, as the intervention induced a considerable fraction of the affected population to select into SP, i.e. acquire insurance coverage, one would expect treatment assignment itself to have an impact on the demand for preventive care—if there is a causal effect.

The intent-to-treat (ITT) effect, denoted by ϕ_s in equation (3), captures this simple intuition.

(3) $$y_{s,i,T} = \mu + \phi_s TREATMENT_{i,T} + \boldsymbol{Q}'_{i,T-1}\boldsymbol{\psi} + \nu_{i,T}.$$

Here s indexes the set of preventive activities listed above, and $y_{s,i,T}$ is an indicator variable for whether individual i engages in s. Therefore, each ϕ_s indicates by how much the treatment group mean changes

[22] I am grateful to Amy Finkelstein for suggesting this approach. Note that it is very similar to the one used by Kling et al. (2007) in evaluating the Moving to Opportunity (MTO) Experiment. Kling et al. (2007) first standardize each outcome variable and then regress the average standardized outcome on treatment assignment. For a discussion of this approach and the one used in this paper see the appendix to Kling et al. (2007).

relative to that of the control group; and the standardized treatment effect expresses the average change over all measures in standard deviation units. Naturally, negative values point towards ex ante moral hazard.

In addition to the effect of treatment on the aggregate demand for preventive care, the causal impact of insurance coverage itself is of interest as well. Using an individual's experimental assignment as instrumental variable (IV), this effect can be estimated by two-stage least squares (2SLS)—treating whether an agent is insured as endogenous and the variables included in $\boldsymbol{Q}_{i,T-1}$ as exogenous. The particular model is given by

$$y_{s,i,T} = \kappa + \pi_s h_{i,T} + \boldsymbol{Q}'_{i,T-1}\boldsymbol{\varsigma} + \eta_{i,T} \tag{4}$$

and the corresponding first stage

$$h_{i,T} = \tau + \delta TREATMENT_{i,T} + \boldsymbol{Q}'_{i,T-1}\boldsymbol{\chi} + \upsilon_{i,T}. \tag{5}$$

All symbols are as defined above.

If there is heterogeneity in the effect of coverage on the demand for preventive services, as seems likely, then $\hat{\pi}_s$ should be interpreted as local average treatment effect (LATE), i.e. as the change in utilization for individuals who enroll in SP, but would have been uninsured had it not been rolled out in their area (Imbens and Angrist 1994).[23]

Table 5 presents estimates of the relationship between insurance coverage and utilization of preventive care. The top row displays average standardized effects. Estimates for each individual measure of self-protection are shown below. The left column contains naïve ordinary least squares (OLS) coefficients, whereas the middle and right column correspond to ITT and LATE estimates (i.e. ϕ and π), respectively.[24]

Strikingly, the standardized effect estimated by OLS is positive, statistically significant, and relatively large. On average, insured individuals demand more preventive care. By contrast, the average standardized ITT and LATE coefficients are negative, statistically (marginally) significant ($p = .050$ and $p = .056$, respectively), and in the latter case about the same magnitude as the one estimated by OLS. That is, the causal effect of insurance goes in the opposite direction. Health insurance coverage induced individuals in the Seguro Popular Experiment to engage in *less* self-protection, indicating the existence of ex ante moral hazard.

Taken at face value the point estimates for individual measures of self-protection support the conclusion drawn above. Every single ITT and LATE estimate is negative, although none of them is

[23] Since the endogenous variable is health insurance status and there are "always takers" in the terminology of Angrist, Imbens, and Rubin (1996), $\hat{\pi}_s$ should not be interpreted as the effect of treatment-on-the-treated (TOT).
[24] When appropriate the sample has been limited to females.

statistically significant by itself. However, compared to dependent variable sample means between 5.6% and 29.2% the point estimates (ranging from .8 to 8.1 percentage points) are economically very large.

Since insurance coverage reduces the full-cost of obtaining preventive medical care, these estimates likely understate the true extent of ex ante moral hazard. That is, holding prices equal, one might expect self-protection to decline even more.[25]

Readers familiar with the original evaluation of the Seguro Popular Experiment might recall that King et al. (2009a) report no change in the utilization of preventative care. It is therefore useful to point out why this paper draws the opposite conclusion. Comparing the results of Table 3 in King et al. (2009a) with those presented in the lower panel of Table 5 in this paper shows that the ITT estimates for individual outcomes are generally very similar. Those with respect to mammograms and eye exams are almost identical, and none of the other point estimates differs by more than .007 (relative to a standard error of .021). The LATE estimates reported in this paper, however, are slightly larger in magnitude than those in King et al. (2009a). The difference between both sets of estimates ranges from .001 to .023 (compared to standard errors between .016 and .057). These differences narrow by about one half, or even reverse, when disregarding the control variables used in this paper. The remainder of the difference between the estimates in both papers is due to the fat that King et al.'s (2009a) coding of $h_{i,T}$ considers only SP affiliates 'insured', whereas this paper is primarily interested in testing the predictions of economic theory and thus also includes all other insured individuals.

However, it is important to note that the difference in interpretation is not an artifact of discrepancies in individual point estimates. All of King et al.'s (2009a) estimates carry the expected sign, i.e. are negative, and are economically very large—between 6% and 18% of the base level. But they are estimated imprecisely. The difference in interpretation between King et al. (2009a) and this paper arises because this paper properly accounts for the testing of multiple hypotheses, i.e. the multiple inference problem. Aggregating information by calculating ASTEs (or z-scores as in Kling et al. 2007, cf. Table 8) avoids the problem of multiple inference and improves statistical power to detect effects that go in the same direction. This estimation strategy provides the econometrician with enough power to reject the null of no change in utilization, and therefore the null of no ex ante moral hazard (for more precise estimates relying on z-scores see Table 8 in the following subsection).

C. Robustness and Sensitivity Analysis

[25] The data contain three more self-reported variables which could potentially proxy for self-protection, i.e. how frequently the respondent smokes, drinks alcohol, and how often she wears a seatbelt. However, self-reported levels of these activities are implausibly low, and factor analysis reveals that these variables do *not* load on the same factor as all the other proxies used in this paper. It is thus a priori doubtful as to whether these variables are indeed reliable measures of self-protection. Yet, in the interest of full disclosure it ought to be pointed out that the effect of SP on these measures is very small and in no case statistically different from zero.

As noted before, one potential concern with the results presented so far is that they could be due to chance differences between the treatment and control group. In fact, at baseline individuals in the treatment group were more likely to have had a mammogram, and were more frequently already covered by health insurance (cf. Table 1). While the former difference would work against detecting a reduction in preventive care, pre-existing disparities in insurance status are potentially more troubling. Therefore, this section performs a variety of sensitivity checks to demonstrate that the results are qualitatively and quantitatively robust.

Table 6 probes the sensitivity of the findings with respect to selection into health insurance. The left panel displays results from linear probability models, whereas the right panel accounts for the binary nature of the dependent variable by estimating probit specifications. Entries correspond to γ in equation (1), or the associated average marginal effects, for all six pre-period proxies of individuals' 'riskiness'. The rightmost column within each set shows point estimates controlling for the set of baseline controls as well as for health insurance status in the pre-period, and the middle column contains estimates without any covariates. For comparison, the coefficients obtained by controlling for $\boldsymbol{Q}_{i,T-1}$ only, i.e. those analogous to the ones in Tables 3 and 4, are shown in the leftmost column. While there do exist differences, it appears that $\hat{\gamma}$ does not depend overly much on the mode of estimation or the set of covariates. That is estimated average marginal effects are very similar to their OLS counterparts.[26]

Controlling only for $\boldsymbol{Q}_{i,T-1}$, Table 7 shows least squares estimates of γ for different subsamples of the data. With two exceptions all coefficients carry the expected sign, i.e. are positive, when pre-period health expenditures are used to proxy for risk. Although the point estimates are not very precise, most coefficients are of roughly similar magnitude suggesting that the results are not driven by one particular group of individuals. With 19 out of 56 entries being negative when objective health indicators are used to proxy for risk, there is again no clear indication of selection along this dimension—as was the case in Table 4.

The sensitivity of the results with respect to ex ante moral hazard is investigated next. Varying the set of controls and the amount of data used in the estimation the columns on the left in Table 8 display average standardized treatment effects, whereas the columns on the right follow Kling et al.'s (2007) evaluation of the Moving to Opportunity Experiment and use z-scores as indices of self-protection.[27] The upper panel contains estimates from cross-sectional models akin to equations (3) and (4), whereas the

[26] Average marginal effects in Table 6 are calculated following Ai and Norton (2003). Graphical representations of the sample distribution of all estimated marginal effects are available from the author upon request. To get a sense of their dispersion it may be helpful to know that their sample standard deviation is generally fairly small, e.g., .017 for self-rated health, .005 for BMI, and .003 for HbA1c (when using the baseline set of covariates).

[27] The point estimates in Table 8 are similar in magnitude when using factor analysis to calculate indices of self-protection, but they are in most cases not statistically significant, or only marginally so, since missing information does not allow these indices to be calculated for a large number of individuals.

lower panel shows coefficients from panel data models including pre-period observations. Going top to bottom within each panel the set of included controls steadily grows.

All cross-sectional point estimates are reasonably close to one another, and, if anything, increase in magnitude after controlling for individuals' health insurance status and the outcome in the pre-period—independently of whether one calculates ASTEs or z-scores (which are somewhat more precisely estimated than the ASTEs). To a lesser extent this also holds for the entries in the lower panel. After controlling for time specific factors common to all individuals by including period fixed effects, the panel data estimates are very similar to their cross-sectional counterparts. It ought to be noted though that including individual fixed effects reduces the ITT and LATE point estimates somewhat, and renders them statistically insignificant. However, they continue to be quite large in an economic sense. It is not entirely clear whether the reduction in the estimates is due to a lack in residual variation—as evidenced by a substantial increase in the standard error—or to true underlying differences between the treatment and control group. The fact that the cross-sectional estimates *increase* after controlling for pre-period outcomes and insurance status suggests that loss in precision might be important.

Average standardized treatment effects for various subsamples of the data are shown in Table 9. In almost all cases do the coefficients carry the expected sign and are of similar magnitude as in Table 5. While large standard errors prevent sharp conclusions, there is some suggestive evidence that ex ante moral hazard is stronger among females.[28]

Broadly summarizing, the sensitivity analysis in this subsection suggests that the main results of this paper are qualitatively quite robust. This may help to alleviate concerns about chance differences between the treatment and control group.

IV. Discussion and Concluding Remarks

Although one ought to be cautious not to overemphasize results from the very specific population studied in this paper, the findings presented above generally confirm the predictions of economic theory and have implications for public policy.

As has been shown by Fang et al. (2008) and Finkelstein and McGarry (2006) for different markets, despite a negative correlation between agents' 'riskiness' and their choice of insurance coverage, *ceteris paribus* individuals do sort based on risk. The crucial question for public policy, however, is whether agents are *on average* adversely or advantageously selected. This paper does not find a source of offsetting advantageous selection; and there remains the possibility that what appears as advantageous selection in the raw data may be due to reverse causality. That is, 'low risk' agents may be more likely to

[28] The larger point estimates for females are not an artifact from changing the set of outcomes. The coefficients are very similar when restricting attention to Flu Shot and Eye Exam only.

work in the formal sector of the Mexican economy, and, therefore, be required to have health insurance. While an insurance mandate for formal sector workers limits the possible extent of selection, this does not imply that risk based sorting will not be important when it comes to voluntary programs such as SP. On one hand this may be welcome—recall that SP was extended to eliminate catastrophic health expenditures among the poor. On the other hand, however, if selection is quantitatively important it needs to be taken into account when assessing the costs (and benefits) of such programs, in particular in cases in which these are designed to be self-sustaining.

The results presented in this paper also point to the existence of ex ante moral hazard, which manifests itself through a reduction in the utilization of preventative care. However, this finding rests on two important assumptions: (i) changes in utilization reflect only demand side changes as opposed to supply side constraints, and (ii) previous levels of preventive care were not inefficiently high.

Given the short evaluation horizon in the SP experiment of only 10 months and the fact that building new medical facilities and accrediting new staff took likely much longer than that, the supply of medical care should been relatively stable. If anything it would appear that the inflow of additional resources caused the production possibility frontier to shift outwards, which implies more and higher quality options for patients in the treatment group compared to those in the control group. If this in indeed the case, then the results in Table 5 will likely not be due to supply side constraints (as these should be less severe in the treatment group), and may even understate the true extent of moral hazard.

Although it is notoriously difficult to determine the optimal level of preventive care, it seems implausible that the mostly rural and overwhelmingly poor population in the SP Experiment received excessive care. In fact, the summary statistics in Table 1 show rather low levels of utilization. For instance, only 5.6% of women received a mammogram, and only 18.5% of individuals had a flu shot.

Taking the results with respect to moral hazard at face value, they have potentially important policy unlikely. In the presence of ex ante moral hazard the welfare consequences of health insurance mandates are theoretically ambiguous, even if agents are adversely selected. Furthermore, optimal insurance contracts will not only trade off risk sharing and ex post moral hazard through co-payments, but would also need to provide appropriate incentives to engage in self-protection. Naturally, the effectiveness of the latter depends crucially on the ability of the insurer to observe precautionary activities. But at the same time, self-protection shares many features with the classic multi-tasking problem (Holmström and Milgrom 1991). This means that the optimal incentive scheme is far from obvious, as it will also depend on the substitutability of various precautionary activities, whether any of them carry externalities—as, for example, flu shots do—and on how much any reduction in self-protection increases ex post cost. Nevertheless, it seems plausible that optimal insurance contracts would either incentivize preventative care through lower co-payments, or even penalties. The exact nature of the optimal incentive provision, however, is left for future research.

* * *

Ever since the seminal contributions of Arrow (1963), Akerlof (1970) as well as Rothschild and Stiglitz (1976) have economists been keenly aware of the possibility of market failure due to asymmetric information. Yet, empirical evidence in favor of these models has been mixed and often not been able to separate moral hazard from selection.

Using data from a large-scale randomized field experiment in Mexico, this paper exploits experimental variation in the availability of health insurance to document systematic sorting based on risk, holding all else equal. Moreover, the evidence presented above suggests the existence of non-negligible ex ante moral hazard; thereby adding to the sparse empirical evidence on the impact of insurance coverage on self-protection.

A. Technical Appendix

CLAIM: Consider the data generating process

$$h_i = \alpha_0 + \rho_0 X_i + \beta_0 Z_i + \gamma_0 Z_i X_i + \epsilon_i.$$

If Z_i is independently distributed of (X_i, ε_i) and the usual full rank condition is satisfied, then the ordinary least squares estimate of γ_0 is well identified. That is, $plim\ \hat{\gamma}_{OLS} = \gamma_0$.

PROOF: The Frisch-Waugh Theorem (Frisch and Waugh 1933) implies that

$$\text{(A.1)} \quad plim\ \hat{\gamma}_{OLS} = \gamma_0 + \frac{Cov(\widetilde{Z_\iota X_\iota}, \varepsilon_i)}{Var(\widetilde{Z_\iota X_\iota})},$$

where $\widetilde{Z_\iota X_\iota}$ denotes the residual from projecting $Z_i X_i$ onto the vector $[1 \quad X_i \quad Z_i]$. With (A.1) in hand, it suffices to show that $Cov\left(\widetilde{Z_\iota X_\iota}, \varepsilon_i\right) = 0$.

From the definition of $\widetilde{Z_\iota X_\iota}$ and using the Frisch-Waugh Theorem again one obtains:

$$\begin{aligned} Cov\left(\widetilde{Z_\iota X_\iota}, \varepsilon_i\right) &= Cov\left(Z_i X_i - \zeta - \tfrac{Cov(Z_i X_i, \tilde{X}_i)}{Var(\tilde{X}_i)} X_i - \tfrac{Cov(Z_i X_i, \tilde{Z}_i)}{Var(\tilde{Z}_i)} Z_i, \varepsilon_i\right) \\ &= Cov\left(\left(Z_i - \tfrac{Cov(Z_i X_i, \tilde{X}_i)}{Var(\tilde{X}_i)}\right) X_i, \varepsilon_i\right), \end{aligned}$$

where $\zeta = \mathbb{E}[Z_i X_i] - \frac{Cov(Z_i X_i, \tilde{X}_i)}{Var(\tilde{X}_i)} \mathbb{E}[X_i] - \frac{Cov(Z_i X_i, \tilde{Z}_i)}{Var(\tilde{Z}_i)} \mathbb{E}[Z_i]$, and $\tilde{X}_i$ $(\tilde{Z}_i)$ corresponds to the residual form projecting X_i (Z_i) onto Z_i (X_i) and a constant.

By applying the definition of the covariance and the Law of Iterated Expectations it follows that

$$Cov\left(\left(Z_i - \tfrac{Cov(Z_i X_i, \tilde{X}_i)}{Var(\tilde{X}_i)}\right) X_i, \varepsilon_i\right) = \mathbb{E}\left[\left(Z_i - \tfrac{Cov(Z_i X_i, \tilde{X}_i)}{Var(\tilde{X}_i)}\right) X_i \varepsilon_i\right] - \mathbb{E}\left[\left(Z_i - \tfrac{Cov(Z_i X_i, \tilde{X}_i)}{Var(\tilde{X}_i)}\right) X_i\right] \mathbb{E}[\varepsilon_i]$$

$$= \left(\mathbb{E}[Z_i] - \frac{Cov(Z_i X_i, \tilde{X}_i)}{Var(\tilde{X}_i)}\right) \mathbb{E}[X_i \varepsilon_i],$$

since Z_i is independent of (X_i, ε_i) and $\mathbb{E}[\varepsilon_i] = 0$.

Note that as Z_i is independent of X_i, $\widetilde{X}_i$ simply corresponds to the deviation of X_i from its mean. Consequently,

$$\frac{Cov(Z_i X_i, \tilde{X}_i)}{Var(\tilde{X}_i)} = \frac{Cov(Z_i X_i, X_i)}{Var(X_i)} = \frac{\mathbb{E}[Z_i X_i^2] - \mathbb{E}[Z_i X_i]\mathbb{E}[X_i]}{Var(X_i)} = \mathbb{E}[Z_i] \frac{\mathbb{E}[X_i^2] - \mathbb{E}[X_i]^2}{Var(X_i)} = \mathbb{E}[Z_i].$$

This shows that $Cov\left(\widetilde{Z_i X_i}, \varepsilon_i\right) = 0$, as desired. ■

B. Selection on Educational Attainment and Pre-Period Health Expenditures

Given that Fang et al. (2008) report offsetting advantageous selection on education and cognitive ability in the Medigap market, it may be of interest to estimate the causal effect of education on insurance coverage in the Seguro Popular experiment. To this end Table B.1 shows estimates of specification (1) replacing $X_{i,T-1}$ with indicator variables for agents' educational achievement. As in Fang et al. (2008), more educated agents are on average more likely to be insured. The causal effect of education on insurance uptake, however, is *negative* and estimated to be very large—at least in the context of SP.[29]

The bottom panel of Table B.1 investigates selection on pre-period health expenditures. Compared to their counterparts who did not incur any out-of-pocket cost in the month preceding the baseline survey, the treatment effect is 3.6 percentage points larger for individuals with positive medical expenditures.[30] Although the least squares coefficient is non-negligible in an economic sense, it is not statistically significant. When controlling for pre-period health-insurance status, however, the estimate increases to 4.1 percentage points and becomes marginally significant (not reported). Hence, there is some suggestive evidence that agents sort into SP based on pre-period health expenditures.

References

Abbring, Jaap H., Pierre-André Chiappori, and Tibor Zavaldi (2008). "Better Safe than Sorry? Ex Ante and Ex Post Moral Hazard in Dynamic Insurance Data," CentER Discussion Paper No. 2008-77.

Abbring, Jaap H., Pierre-André Chiappori, and Jean Pinquet (2003). "Moral Hazard and Dynamic Insurance Data," *Journal of the European Economic Association, 1*, 767-820.

Ai, Chungron, and Edward C. Norton (2003). "Interaction Terms in Logit and Probit Models," *Economics Letters, 80*, 123-129.

[29] This result is not an artifact of highly educated individuals being more likely to have insurance in the pre-period. The point estimates are almost identical when restricting attention to previously uninsured agents.

[30] On the baseline survey approximately 75% of individuals report zero medical expenditures during the last month.

Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin (1996). "Identification of Causal Effects Using Instrumental Variables," *Journal of the American Statistical Association, 91*, 444-455.

Arrow, Kenneth J. (1963). "Uncertainty and the Welfare Economics of Medical Care," *American Economic Review, 53*, 941-973.

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

Barros, Rodrigo (2008). "Wealthier But Not Much Healthier: Effects of a Health Insurance Program for the Poor in Mexico," Unpublished Manuscript. Stanford University.

Bleich, Sara N., David M. Cutler, Alyce S. Adams, Rafael Lozano, and Christopher J. Murray (2007). "Impact of Insurance and Supply of Health Professionals on Coverage and Treatment for Hypertension in Mexico: Population Based Study," *British Medical Journal, 335*, 875.

Card, David, Carlos Dobkin, and Nicole Maestas (2008). "The Impact of Nearly Universal Insurance Coverage on Health Care: Evidence from Medicare," *American Economic Review 98*, 2242-2258

Cawley, John, and Tomas Philipson (1999). "An Empirical Examination of Information Barriers to Trade in Insurance," *American Economic Review, 89*, 827-846.

Chiappori, Pierre-André, and Bernard Salanié (2000). "Testing for Asymmetric Information in Insurance Markets," *Journal of Political Economy, 108*, 56-78.

Chiappori, Pierre-André, and Bernard Salanié (2003). "Testing Contract Theory: A Survey of some Recent Work," (pp. 115-149) in Mathias Dewatripont, Lars P. Hansen, and Stephen J. Turnovsky, eds., Advances in Economics and Econometrics, Cambridge, UK: Cambridge University Press.

Chiappori, Pierre-André, Bruno Julien, Bernard Salanié, and François Salanié (2006). "Asymmetric Information in Insurance: General Testable Implication," *RAND Journal of Economics, 37*, 783-798.

Cutler, David M., and Richard J. Zeckhauser (2000). "The Anatomy of Health Insurance," (pp. 563-643) in Anthony J. Culyer and Joseph P. Newhouse, eds., *Handbook of Health Economics*, Amsterdam: Elsevier.

de Meza, David, and David C. Webb (2001). "Advantageous Selection in Insurance Markets," *RAND Journal of Economics*, 32, 249-262.

Ehrlich, Isaac, and Gary S. Becker (1972). "Market Insurance, Self-Insurance, and Self-Protection," *Journal of Political Economy, 80*, 623-648.

Einav, Liran, Amy Finkelstein, and Paul Schrimpf (2010). "Optimal Mandates and The Welfare Cost of Asymmetric Information: Evidence from the U.K. Annuity Market," forthcoming in *Econometrica*.

Fang, Hanming, Michael P. Keane, and Dan Silverman (2008). "Sources of Advantageous Selection: Evidence from the Medigap Insurance Market," *Journal of Political Economy, 116*, 303-350.

Finkelstein, Amy, and Kathleen McGarry (2006). "Multiple Dimensions of Private Information: Evidence from the Long-Term Care Insurance Market," *American Economic Review, 96*, 938-958.

Finkelstein, Amy, and James Poterba (2002). "Selection Effects in the Market for Individual Annuities: New Evidence from the United Kingdom," *Economic Journal, 112*, 28-50.

Finkelstein, Amy, and James Poterba (2004). "Adverse Selection in Insurance Markets: Policyholder Evidence from the U.K. Annuity Market," *Journal of Political Economy, 112*, 183-208.

Frenk, Julio, Eduardo González-Pier, Octavio Gómez-Dentéz, Miguel A. Lezana, and Felicia M. Knaul (2006). "Comprehensive Reform to Improve Health System Performance in Mexico," *Lancet, 368*, 1524-1534.

Frisch, Ragnar, and Frederick V. Waugh (1933). "Partial Time Regressions as Compared with Individual Trends," *Econometrica, 1*, 387-401.

Gakido, Emmanuela, Rafael Lozano, Eduardo González-Pier, Jesse Abbott-Klafter, Jeremy T. Barofsky, Chloe Bryson-Cahn, Dennis M. Feehan, Dianna K. Lee, Hector Hernández-Llamas, Christopher and J.L. Murray (2006). "Assessing the Effect of the 2001-06 Mexican Health Reform: An Interim Report Card," *Lancet, 368*, 1920-1935.

Holmström, Bengt, and Paul Milgrom (1991). "Multi-Task Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design," *Journal of Law, Economics, and Organization 7*, 24-52.

Homedes, Nuria, and Antonio Ugalde (2009). "Twenty-Five Years of Convoluted Health Reforms in Mexico," *PLoS Medicine, 6*, e1000124.

Horowitz, Joel L. (2001). "The Bootstrap in Econometrics," (pp. 3159-3228) in James J. Heckman and Edward E. Leamer, eds., *Handbook of Econometrics*, vol. 5, Amsterdam: Elsevier.

Imbens, Guido W. and Joshua D. Angrist (1994). "Identification and Estimation of Local Average Treatment Effects," *Econometrica, 62*, 467-475.

King, Gary, Emmanuela Gakidou, Nirmala Ravishankar, Ryan T. Moore, Jason Lakin, Manett Vargas, Martha María Téllez-Rojo, Juan Eugenio Hernández-Ávila, Mauricio Hernández-Ávila, and Hector Hernández Llamas (2007). "A 'Politically Robust' Experimental Design for Public Policy Evaluation, with Application to the Mexican Universal Health Insurance Program," *Journal of Policy Analysis and Management, 26*, 479-506.

King, Gary, Emmanuela Gakidou, Kosuke Imai, Jason Lakin, Ryan T. Moore, Clayton Nall, Nirmala Ravishankar, Manett Vargas, Martha María Téllez-Rojo, Juan Eugenio Hernández-Ávila, Mauricio Hernández-Ávila, and Hector Hernández Llamas (2009a). "Public Policy for the Poor? A Randomised Assessment of the Mexican Universal Health Insurance Programme," *Lancet, 373*, 1447-1454.

King, Gary, Emmanuela Gakidou, Kosuke Imai, Jason Lakin, Ryan T. Moore, Clayton Nall, Nirmala Ravishankar, Manett Vargas, Martha María Téllez-Rojo, Juan Eugenio Hernández-Ávila, Mauricio Hernández-Ávila, and Hector Hernández Llamas (2009b). "Replication data for: Public Policy for the Poor? A Randomised Assessment of the Mexican Universal Health Insurance Programme," available online at http://hdl.handle.net/1902.1/11044

Klick, Jonathan, Thomas Stratmann (2007). "Diabetes Treatment and Moral Hazard," *Journal of Law and Economics, 50*, 519-538.

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz (2007). "Experimental Analysis of Neighborhood Effects," *Econometrica, 75*, 83-119.

Knox, Melissa A. (2008). "Health Insurance for All: An Evaluation of Mexico's Seguro Popular Program." Unpublished Manuscript. University of California, Berkeley.

Lillard, Lee A., Willard G. Manning, Christine E. Peterson, Nicole Lurie, George A. Goldberg, and Charles E. Phelps (1986). *Preventive Medical Care: Standards, Usage, and Efficacy*. Santa Monica, CA: RAND Corporation.

National Institutes of Health (1998). *Clinical Guidelines on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults.* NIH Publication No. 98-4083

National Institutes of Health (2002). *Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III).* NIH Publication No. 02-5215

National Institutes of Health (2004). *The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure*. NIH Publication No. 04-5230.

Newhouse, Joseph P., and the Insurance Experiment Group (1993). *Free For All? Lessons from the RAND Health Insurance Experiment*. Cambridge, MA: Harvard University Press.

Organisation for Economic Co-operation and Development (2005). *OECD Reviews of Health Systems: Mexico*. Paris: OECD Publications.

Pauly, Mark V. (1974) "Overinsurance and Public Provision of Insurance: The Roles of Moral Hazard and Adverse Selection," *Quarterly Journal of Economics, 88*, 44-62.

Rothschild, Micheal, and Joseph Stiglitz (1976). "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information," *Quarterly Journal of Economics, 90*, 629-649.

Sosa-Rubi, Sandra G., Omar Galárraga, and Jeffrey E. Harris (2009). "Heterogeneous Impact of the 'Seguro Popular' Program on the Utilization of Obstetrical Services in Mexico, 2001-2006: A Multinomial Probit Model with a Discrete Endogenous Variable," *Journal of Health Economics, 28*, 20-34.

World Health Organization (2000). *Obesity: Preventing and Managing the Global Epidemic*. Technical Report Series No 894. Geneva: World Health Organization.

Zeckhauser, Richard (1970). "Medical Insurance: A Case Study of the Tradeoff between Risk Spreading and Appropriate Incentives," *Journal of Economic Theory, 2*, 10-26.

Zweifel, Peter, and Willard G. Manning (2000). "Moral Hazard and Consumer Incentives in Health Care," (pp. 409-459) in Anthony J. Culyer and Joseph P. Newhouse, eds., *Handbook of Health Economics*, Amsterdam: Elsevier.

Data Appendix to
"Moral Hazard and Selection Among the Poor: Evidence from a Randomized Experiment"

All data used in this paper stem from the Seguro Popular Assessment by King et al. (2007, 2009a, 2009b). This appendix provides a description of the key variables used in the analysis. For a detailed description of the experimental design beyond that in Section II.B the interested reader should consult King et al. (2007, 2009a).

A. Outcome Variables

Health Insurance is an indicator variable equal to one if (on the first section of the household survey) the survey respondent is reported to be covered by any health insurance, and zero otherwise. The variable is available on for pre- and post-period.

BMI is defined as the respondent's weight (in kilograms) over her squared height (in meters). Weight and height were both elicited together with other socio-demographic characteristics during the pre- and post-period household surveys. A person's BMI is said to be outside its normal range if it exceeds 25 kg/m^2 or is lower than 18 kg/m^2 (cf. NIH 1998, WHO 2000).

Blood Pressure is the mean of two systolic blood pressure measurements conducted at the beginning and at the end of the household survey, respectively. A person's systolic blood pressure is defined to be outside its normal range if it exceeds 120 mmHg (cf. NIH 2004). The variable is available for the pre- and post-period.

Cholesterol is a respondent's cholesterol level in mg/dL, as measured at the end of the household survey. A person's cholesterol is defined to be outside its normal range if it exceeds 200 mg/dL (cf. NIH 2002). The variable is available for the pre- and post-period.

HbA1c is a respondent's level of glycated hemoglobin (in percent), as measured at the end of the household survey. HbA1c is an indicator of blood sugar levels over the last couple of months, and thus often used in the diagnosis of diabetes. A person's HbA1c level is defined to be outside its normal range if it exceeds 6% (cf. NIH 1998). The variable is available for the pre- and post-period.

Flu Shot is an indicator variable equal to one if (on the household survey) the respondent reports to have been vaccinated against the flu within the last 12 months, and zero otherwise. The variable is available for the pre- and post-period.

Pelvic Exam is an indicator variable equal to one if (on the household survey) the respondent reports to have had a pelvic examination within the last 12 months, and zero otherwise. The variable is available for the pre- and post-period, but only for females.

Pap Smear is an indicator variable equal to one if (on the household survey) the respondent reports to have had a pap smear test within the last 12 months, and zero otherwise. The variable is available for the pre- and post-period, but only for females.

Mammogram is an indicator variable equal to one if (on the household survey) the respondent reports to have had a mammography within the last 12 months, and zero otherwise. The variable is available for the pre- and post-period, but only for females.

Eye Exam is an indicator variable equal to one if (on the household survey) the respondent reports to have had her eyes examined by a health care professional within the last 12 months, and zero otherwise. The variable is available for the pre- and post-period.

B. Independent Variables

Self-Assessed Health corresponds to a respondent's answer to the question "In general, how would you rate your health today?" The set of possible answers consisted of: "very good", "good", "fair", "bad", "very bad". This paper combines the first two and the last two answer choices, respectively, to form a set of three indicator variables. These variables are available for the pre- and post-period.

Health Expenditures is defined as the household's expenditures on "health care costs, excluding travel expenses related to seeking health care and any reimbursement of health insurance" (in pesos) during the last month. The variable is available on for pre- and post-period.

Gender is an indicator variable equal to one if the respondent is female, and zero otherwise.

Age is the respondent's age (in years), as indicated on the household survey.

Educational Achievement encompasses five indicator variables capturing the respondent's formal education, as reported on the post-period household survey. The categories considered in this paper are: 'less than primary school', 'primary school', 'completed middle school', 'completed high school', and 'at least some post-secondary education'. The first two serve as omitted category in all regressions that include covariates.

Total Household Expenditures is defined as the household's total expenditures (in pesos) during the last month. This variable is available for the pre- and post-period.

High Asset Holdings is defined as an indicator variable equal to one if the household has at least half of the following 'items', and zero otherwise: a cement or tile floor, electricity, a washing machine, a gas stove, a refrigerator, a phone, a TV set, a computer, or a second home. This variable has been created by King et al. (2009a), and is available on the pre- and post-period period.

Number of Doctors gives the number of doctors in a particular health cluster prior to the intervention.

Number of Nurses gives the number of nurses in a particular health cluster prior to the intervention.

Oportunidades is an indicator variable equal to one if the respondent's household participated in the Opotunidades anti-poverty program, and zero otherwise. The variable is available on the pre- and post-period survey. Opotunidades has formerly been known as Progresa.

Urban is an indicator variable equal to one if the respondent's household is located in an urban health cluster, and zero otherwise.

Indigenous is an indicator variable equal to one if (on the pre-period household survey) the respondent reports speaking one of Mexico's indigenous languages or dialects, and zero otherwise.

References

King, Gary, Emmanuela Gakidou, Nirmala Ravishankar, Ryan T. Moore, Jason Lakin, Manett Vargas, Martha María Téllez-Rojo, Juan Eugenio Hernández-Ávila, Mauricio Hernández-Ávila, and Hector Hernández Llamas (2007). "A 'Politically Robust' Experimental Design for Public Policy Evaluation, with Application to the Mexican Universal Health Insurance Program," *Journal of Policy Analysis and Management, 26*, 479-506.

King, Gary, Emmanuela Gakidou, Kosuke Imai, Jason Lakin, Ryan T. Moore, Clayton Nall, Nirmala Ravishankar, Manett Vargas, Martha María Téllez-Rojo, Juan Eugenio Hernández-Ávila, Mauricio Hernández-Ávila, and Hector Hernández Llamas (2009a). "Public Policy for the Poor? A Randomised Assessment of the Mexican Universal Health Insurance Programme," *Lancet, 373*, 1447-1454.

King, Gary, Emmanuela Gakidou, Kosuke Imai, Jason Lakin, Ryan T. Moore, Clayton Nall, Nirmala Ravishankar, Manett Vargas, Martha María Téllez-Rojo, Juan Eugenio Hernández-Ávila, Mauricio Hernández-Ávila, and Hector Hernández Llamas (2009b). "Replication data for: Public Policy for the Poor? A Randomised Assessment of the Mexican Universal Health Insurance Programme," available online at http://hdl.handle.net/1902.1/11044

National Institutes of Health (1998). *Clinical Guidelines on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults.* NIH Publication No. 98-4083.

National Institutes of Health (2002). *Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III).* NIH Publication No. 02-5215.

National Institutes of Health (2004). *The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure*. NIH Publication No. 04-5230.

World Health Organization (2000). *Obesity: Preventing and Managing the Global Epidemic*. Technical Report Series No 894. Geneva: World Health Organization.

Figure 1: Location of Treatment and Control Clusters

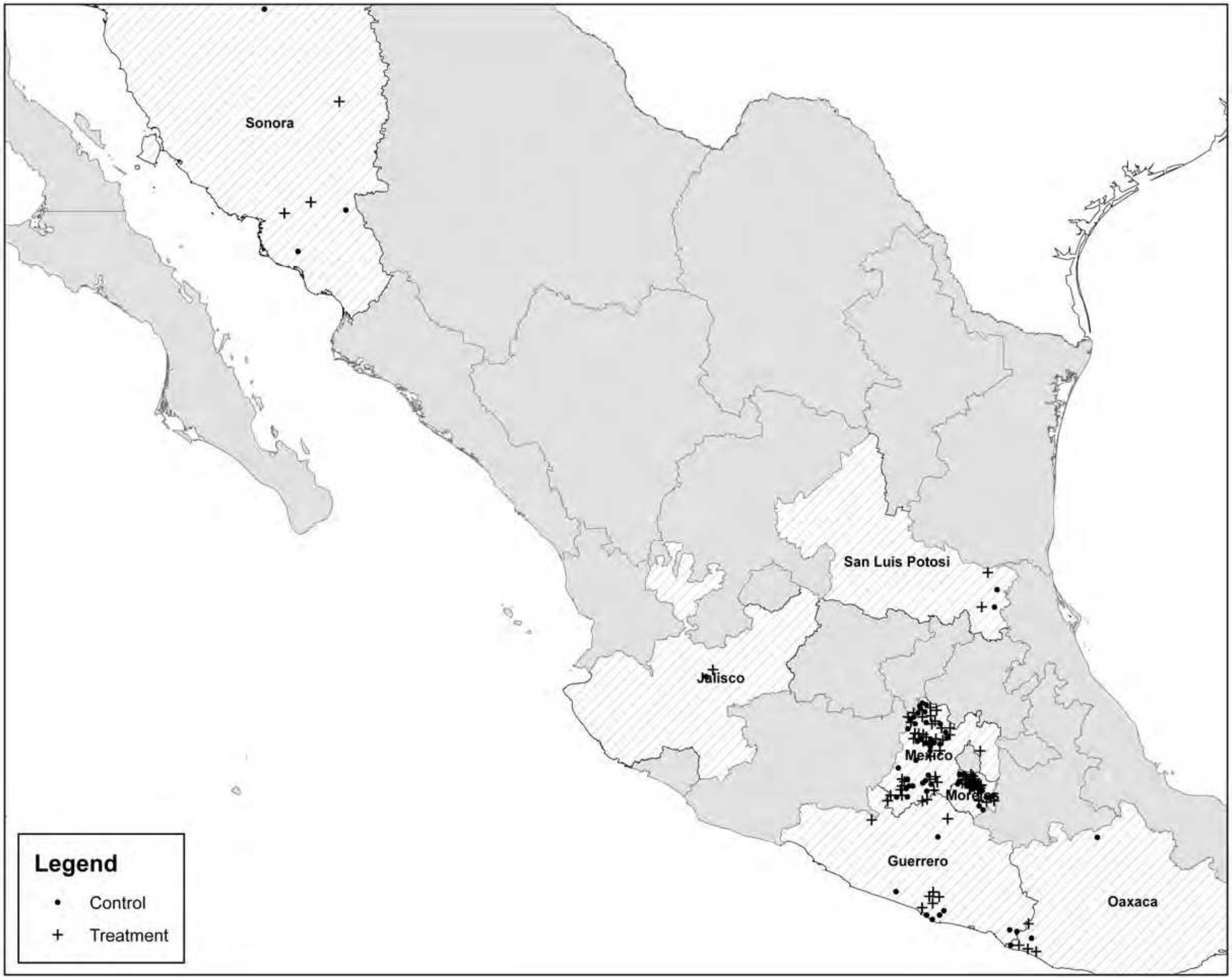

Table 1: Pre-Period Summary Statistics

	Full Sample		Treatment Group		Control Group		*p*-value
	Mean	SD	Mean	SD	Mean	SD	Treatment = Control
Demographics:							
Female	.620	(.485)	.627	(.483)	.613	(.487)	.21
Age	42.24	(16.75)	42.36	(16.72)	42.13	(16.78)	.69
Indigenous	.060	(.236)	.064	(.244)	.057	(.231)	.82
Urban	.106	(.308)	.109	(.312)	.103	(.303)	.92
Oportunidades	.451	(.498)	.458	(.498)	.445	(.497)	.81
Education:							
Less than Primary School	.446	(.497)	.442	(.497)	.451	(.498)	.76
Primary School School	.246	(.430)	.239	(.427)	.252	(.434)	.27
Incomplete Secondary School	.193	(.395)	.195	(.396)	.191	(.393)	.76
Seondary School	.049	(.216)	.051	(.219)	.048	(.213)	.76
Post-Secondary Schooling	.066	(.248)	.073	(.261)	.059	(.235)	.33
Household Characteristics:							
Household Size	4.085	(2.201)	4.132	(2.232)	4.039	(2.169)	.33
High Asset Household	.546	(.498)	.558	(.497)	.535	(.499)	.63
Annualized Household Expenditures	36,159	(46,047)	37,035	(45,803)	35,283	(46,986)	.54
Annualized Expenditures on Health Care	1,483	(6,303)	1,492	(6,115)	1,473	(6,487)	.92
Health Insurance Status:							
Any Health Insurance	.218	(.413)	.255	(.436)	.181	(.385)	.05
Obligatory Health Insurance	.154	(.361)	.167	(.373)	.140	(.347)	.35
Voluntary Health Insurance	.070	(.255)	.093	(.290)	.047	(.213)	.07
Self-Rated Health:							
Very Good or Good	.573	(.495)	.583	(.493)	.564	(.496)	.35
Fair	.371	(.483)	.364	(.481)	.379	(.485)	.43
Bad or Very Bad	.055	(.228)	.053	(.223)	.058	(.233)	.32
Objective Health Measures:							
BMI (kg/m^2)	26.03	(5.18)	26.02	(5.04)	26.03	(5.30)	.96
Systolic Blood Pressure (mmHg)	125.4	(18.9)	125.6	(19.3)	125.2	(18.6)	.49
Cholesterol (mg/dL)	176.1	(27.6)	174.9	(27.3)	177.3	(27.9)	.25
HbA1c (%)	6.062	(1.328)	6.100	(1.346)	6.023	(1.307)	.25
Preventive Care:							
Flu Shot	.185	(.388)	.174	(.379)	.196	(.397)	.21
Pelvic Exam	.222	(.416)	.212	(.409)	.233	(.423)	.34
Pap Smear	.292	(.455)	.289	(.453)	.295	(.456)	.80
Mammogram	.056	(.230)	.064	(.244)	.048	(.214)	.04
Eye Exam	.108	(.310)	.110	(.313)	.106	(.307)	.65
Other Medical Utilization Within Last Year:							
Outpatient Care	.518	(.500)	.513	(.500)	.522	(.499)	.67
Hospitalized	.069	(.254)	.070	(.255)	.069	(.254)	.95
Medical Staff:							
Number of Doctors	1.645	(1.719)	1.597	(1.319)	1.721	(2.040)	.67
Number of Nurses	1.581	(2.306)	1.423	(1.360)	1.739	(2.956)	.51
Number of Observations	32,426		16,214		16,212		
Number of Clusters	100		50		50		

Notes: Entries are means and standard deviations of pre-period data for those individuals with non-missing information on insurance status. The right column displays *p*-values from testing the equality of means in the treatment and control group. See the Data Appendix for a precise definition of each variable.

Table 2: Health Insurance Status in the Seguro Popular Experiment

	Control Group		Treatment Group		
	Pre-Period	Post-Period	Pre-Period	Post-Period	Diff-in-Diff
Any Health Insurance:	.181	.212	.255	.561	.276
	(.023)	(.026)	(.029)	(.026)	(.037)
Obligatory Health Insurance:	.140	.130	.167	.156	-.001
	(.018)	(.017)	(.022)	(.020)	(.009)
IMSS	.097	.089	.107	.100	.001
	(.014)	(.012)	(.017)	(.015)	(.007)
IMSSTE	.025	.025	.032	.031	-.001
	(.005)	(.005)	(.006)	(.006)	(.003)
PEMEX	.000	.000	.000	.001	.000
	(.000)	(.000)	(.000)	(.000)	(.000)
SEDENA	.002	.001	.001	.002	.001
	(.001)	(.000)	(.000)	(.001)	(.001)
MARINA	.000	.000	.002	.001	-.000
	(.000)	(.000)	(.001)	(.000)	(.001)
Other	.015	.013	.025	.022	-.001
	(.003)	(.003)	(.006)	(.005)	(.002)
Voluntary Health Insurance:	.047	.087	.093	.413	.280
	(.011)	(.021)	(.022)	(.035)	(.035)
Seguro Popular	.030	.053	.077	.393	.292
	(.011)	(.015)	(.021)	(.035)	(.032)
Private	.005	.003	.004	.002	.002
	(.002)	(.001)	(.001)	(.001)	(.001)
IMSS voluntarily	.005	.003	.003	.003	.002
	(.001)	(.001)	(.001)	(.001)	(.001)
Other	.008	.029	.009	.014	-.017
	(.002)	(.012)	(.005)	(.004)	(.014)

Notes: Entries denote the fraction of individuals who report being affiliated with different health insurance options, by treatment assignment and period. Standard errors are clustered on the level of randomization and reported in parentheses. The right column displays difference-in-differences estimates and the corresponding standard errors. Note that the categories are not mutually exclusive.

Table 3: Selection on Self-Rated Health

	Post-Period Health Insurance Status	
Self-Rated Health in Pre-Period:		
In Fair Health	-.013	-.040
	(.009)	(.013)
In Bad or Very Bad	-.041	-.075
	(.015)	(.022)
Treatment	.345	.320
	(.035)	(.037)
Treatment × In Fair Health		.055
		(.019)
Treatment × In Bad or Very Bad Health		.069
		(.031)
Baseline Controls	Yes	Yes
R-Squared	.173	.174
Number of Observations	28,247	28,247

Notes: Entries are coefficients and standard errors from estimating equation (1) by ordinary least squares with individuals' subjective health assessments in the pre-period proxying for risk. The omitted category is 'good or very good health'. Heteroskedasticity robust standard errors are clustered on the level of randomization and reported in parentheses. In addition to the variables shown in the table, indicator variables for missing values on each covariate are also included in the regressions. See the Data Appendix for the precise definition and source of each variable.

Table 4: Selection on Objective Health Measures

A. Pre-Period BMI

	Post-Period Health Insurance Status	
BMI Within Normal Range	-.017	-.020
	(.008)	(.009)
Treatment	.336	.334
	(.034)	(.033)
Treatment × BMI Within Normal Range		.006
		(.016)
Baseline Controls	Yes	Yes
R-Squared	.168	.168
Number of Observations	19,292	19,292

B. Pre-Period Systolic Blood Pressure

	Post-Period Health Insurance Status	
Blood Pressure Within Normal Range	-.010	-.018
	(.007)	(.010)
Treatment	.346	.339
	(.035)	(.035)
Treatment × Blood Pressure Within Normal Range		.016
		(.014)
Baseline Controls	Yes	Yes
R-Squared	.173	.174
Number of Observations	28,092	28,092

C. Pre-Period Cholesterol

	Post-Period Health Insurance Status	
Cholesterol Within Normal Range	-.005	.003
	(.016)	(.022)
Treatment	.358	.373
	(.034)	(.043)
Treatment × Cholesterol Within Normal Range		-.018
		(.032)
Baseline Controls	Yes	Yes
R-Squared	.176	.176
Number of Observations	23,217	23,217

D. Pre-Period HbA1c

	Post-Period Health Insurance Status	
HbA1c Within Normal Range	.012	-.016
	(.014)	(.020)
Treatment	.365	.331
	(.035)	(.040)
Treatment × HbA1c Within Normal Range		.053
		(.029)
Baseline Controls	Yes	Yes
R-Squared	.185	.185
Number of Observations	13,398	13,398

Notes: Entries are coefficients and standard errors from estimating equation (1) by ordinary least squares with different anthropometric measures of health proxying for risk. Heteroskedasticity robust standard errors are clustered on the level of randomization and reported in parentheses. In addition to the variables shown in the table, indicator variables for missing values on each covariate are also included in the regressions. See the Data Appendix for the precise definition and source of each variable.

Table 5: The Impact of Insurance Coverage on Utilization of Preventive Care

	Pre-Period	Post-Period	
	OLS	ITT	LATE
Standardized Effect	.145	-.047	-.131
	(.022)	(.024)	(.069)
Witin Last Year:			
Flu Shot	.073	-.023	-.065
	(.012)	(.019)	(.056)
Pelvic Exam	.040	-.020	-.054
	(.014)	(.021)	(.057)
Pap Smear	.070	-.030	-.081
	(.017)	(.020)	(.054)
Mammogram	.033	-.009	-.024
	(.007)	(.006)	(.016)
Eye Exam	.045	-.008	-.024
	(.008)	(.008)	(.024)

Notes: Entries are least squares estimates of the effect of insurance coverage on the utilization of preventive care. The left column displays naïve ordinary least squares estimates, whereas the columns on the right contain ITT and LATE estimates corresponding to ϕ and π in equations (3) and (4), respectively. Average standardized treatment effects are reported in the top row, and estimates for each individual preventive activity are shown below. Bootstrapped standard errors are clustered on the level of randomization and reported in parentheses. All regressions contain the baseline set of covariates as well as indicator variables for missing values on each covariate. See the Data Appendix for the precise definition and source of each variable.

Table 6: Robustness and Sensitivity of Selection Results

	I. Linear Probability Model Coefficient on Interaction Term			*II. Probit* Average Marginal Effect		
Self-Rated Health in Pre-Period:						
In Fair Health	.055	.051	.048	.052	.051	.044
	(.019)	(.019)	(.017)	(.019)	(.019)	(.017)
In Bad or Very Bad Health	.069	.060	.055	.054	.058	.043
	(.031)	(.032)	(.025)	(.029)	(.031)	(.024)
Objective Health Indicator Within Normal Range:						
Pre-Period BMI	.006	.005	.007	.009	.005	.006
	(.016)	(.017)	(.013)	(.016)	(.017)	(.013)
Pre-Period Cholesterol	-.018	-.020	-.018	-.018	-.020	-.018
	(.032)	(.031)	(.025)	(.025)	(.030)	(.024)
Pre-Period Systolic Blood Pressure	.016	.014	.020	.019	.013	.020
	(.014)	(.014)	(.011)	(.014)	(.014)	(.012)
Pre-Period HbA1c	.053	.049	.041	.058	.049	.041
	(.029)	(.028)	(.025)	(.028)	(.027)	(.025)
Baseline Controls	Yes	No	Yes	Yes	No	Yes
Pre-Period Health Insurance Status	No	No	Yes	No	No	Yes

Notes: Entries denote estimates of γ in equation (1), or the corresponding marginal effects, for various proxies of risk and different sets of covariates. Results using linear probability models are shown in the panel on the left, whereas the right panel contains average marginal effects estimated from probit models. The leftmost column within each set contains estimates analogous to those in Tables 3-6, i.e. controlling only for the baseline covariates. The middle column displays coefficients without controlling for any other observables, and the rightmost column controls for health insurance status in the pre-period as well as for the baseline covariates. The respective risk proxy is listed on the left of each row. Standard errors are clustered on the level of randomization and reported in parentheses. All regressions contain indicator variables for missing values on each covariate. See the Data Appendix for the precise definition and source of each variable.

Table 7: Selection by Subsample

Sample	Self-Rated Health: Fair	Self-Rated Health: Bad or Very Bad	Health Indicator Within Normal Range: BMI	Cholesterol	Blood Pressure	HbA1c
Full Sample	.055	.069	.006	-.018	.016	.053
	(.019)	(.031)	(.016)	(.032)	(.014)	(.029)
By Age:						
< 40	.059	.111	-.007	.013	.031	.064
	(.022)	(047)	(.019)	(.035)	(.017)	(.036)
40 to 60	.050	.086	.028	-.048	.012	.068
	(.026)	(.044)	(.030)	(.033)	(.023)	(.035)
> 60	.069	.040	.018	-.020	-.011	.065
	(.035)	(.049)	(.034)	(.052)	(.036)	(.043)
By Gender:						
Females	.090	.055	-.023	-.061	.000	.041
	(.025)	(.051)	(.023)	(.037)	(.019)	(.035)
Males	.032	.061	.028	.006	.007	.056
	(.020)	(.034)	(.020)	(.033)	(.016)	(.031)
By Household Assests:						
High Assets	.054	.088	-.001	-.010	.007	.061
	(.021)	(.043)	(.021)	(.032)	(.018)	(.028)
Low Assets	.051	.025	-.020	-.037	.024	.025
	(.030)	(.041)	(.023)	(.041)	(.018)	(.040)
By Pre-Period Employement Status:						
Employed	.076	.101	-.013	-.021	.019	.037
	(.023)	(.043)	(.020)	(.036)	(.018)	(.033)
Unemployed	.033	.036	.033	-.011	.009	.072
	(.022)	(.035)	(.022)	(.034)	(.018)	(.034)
By Location Type:						
Urban	-.050	-.070	-.008	.073	-.005	.068
	(.044)	(.103)	(.034)	(.055)	(.043)	(.055)
Rural	.051	.064	.003	-.040	.017	.049
	(.020)	(.032)	(.017)	(.033)	(.013)	(.030)
By Number of Doctors in the Area:						
Above Median	.026	.114	-.005	.041	.028	.028
	(.027)	(.048)	(.025)	(.037)	(.024)	(.035)
Below Median	.051	.028	.002	-.032	.012	.045
	(.022)	(.034)	(.019)	(.035)	(.016)	(.034)

Notes: Entries denote ordinary least squares of γ in equation (1) for various proxies of risk and different subsamples of the data. The respective risk proxy is listed on the top of each column, and the subsample is listed on the left of each row. Heteroskedasticity robust standard errors are clustered on the level of randomization and reported in parentheses. All regressions contain the baseline set of controls and indicator variables for missing values on each covariate. See the Data Appendix for the precise definition and source of each variable.

Table 8: Robustness and Sensitivity Analysis of the Impact of Health Insurance on Self-Protection

A. Cross-Sectional Analysis (Post-Period)

	I. ASTE Preventive Care		*II. Z-Scores* Preventive Care	
Covariates	ITT	LATE	ITT	LATE
No Controls	-.038	-.103	-.037	-.106
	(.025)	(.069)	(.017)	(.051)
Baseline Controls	-.047	-.131	-.044	-.128
	(.024)	(.069)	(.016)	(.051)
Baseline Controls, Pre-Period Health Insurance	-.056	-.174	-.051	-.161
	(.024)	(.076)	(.016)	(.056)
Baseline Controls, Pre-Period Health Insurance,	-.050	-.152	-.046	-.145
Outcome in Pre-Period	(.022)	(.071)	(.014)	(.050)

B. Panel Data Analysis (Pre- & Post-Period)

	I. ASTE Preventive Care		*II. Z-Scores* Preventive Care	
Covariates	ITT	LATE	ITT	LATE
No Controls	-.023	-.061	-.034	-.098
	(.018)	(.048)	(.012)	(.038)
Period Fixed Effects	-.038	-.103	-.037	-.106
	(.025)	(.069)	(.017)	(.051)
Period Fixed Effects, Baseline Controls	-.047	-.133	-.044	-.130
	(.024)	(.068)	(.016)	(.051)
Period Fixed Effects, Individual Fixed Effects	-.027	-.095	-.031	-.112
	(.028)	(.109)	(.028)	(.071)

Notes: Entries are ITT and LATE estimates of the effect of insurance coverage on the utilization of preventive care analogous to ϕ and π in equations (3) and (4), respectively. The columns on the left show average standardized treatment effects (cf. equation (2)), and the colums on the right use z-scores as indices of self-protection (cf. Kling et al. 2007). The top panel presents cross-sectional estimates, and the lower one contains coefficients from panel data models using pre- and post-period observations. The set of controls is denoted on the left of each row. Standard errors are clustered on the level of randomization and reported in parentheses. Those for ASTE have been calculated using the block-bootstrap with 1,000 iterations. All regressions contain indicator variables for missing values on each covariate. See the Data Appendix for the precise definition and source of each variable.

Table 9: Average Standardized Treatment Effects by Subsample

Sample	Preventive Care ITT	Preventive Care LATE
Full Sample	-.047	-.131
	(.024)	(.069)
By Age:		
< 40	-.039	-.109
	(.026)	(.070)
40 to 60	-.054	-.149
	(.025)	(.073)
> 60	-.066	-.198
	(.036)	(.122)
By Gender:		
Females	-.053	-.145
	(.024)	(.067)
Males	-.012	-.037
	(.027)	(.098)
By Household Assests:		
High Assets	-.037	-.121
	(.024)	(.081)
Low Assets	-.065	-.149
	(.030)	(.081)
By Pre-Period Employement Status:		
Employed	-.058	-.162
	(.026)	(.077)
Unemployed	-.048	-.129
	(.025)	(.068)
By Location Type:		
Urban	-.052	-1.008
	(.042)	(49.022)
Rural	-.048	-.122
	(.026)	(.068)
By Number of Doctors in the Area:		
Above Median	-.063	-.189
	(.047)	(.143)
Below Median	-.035	-.087
	(.027)	(.072)
By Pre-Period Health:		
In Very Good or Good Health	-.037	-.109
	(.027)	(.078)
In Fair Health	-.057	-.147
	(.025)	(.070)
In Bad or Very Bad Health	-.068	-.164
	(.040)	(.098)

Notes: Entries are standardized average treatments effects of insurance coverage on the utilization of preventive care in various subsamples of the data, i.e. ITT and LATE estimates corresponding to ϕ and π in equations (3) and (4), respectively. The respective subsample is listed on the left of each row. Bootstrapped standard errors are clustered on the level of randomization and reported in parentheses. All regressions contain indicator variables for missing values on each covariate. See the Data Appendix for the precise definition and source of each variable.

Table B.1: Selection on Educational Attainment and Health Expenditures

A. Educational Attainment

	Post-Period Health Insurance Status	
Educational Achievement:		
Completed Middle School	.040	.082
	(.014)	(.019)
Completed High School	.076	.168
	(.023)	(.030)
Post-Secondary Schooling	.215	.354
	(.021)	(.031)
Treatment	.345	.383
	(.035)	(.038)
Treatment × Completed Middle School		-.084
		(.028)
Treatment × Completed High School		-.180
		(.047)
Treatment × Post-Secondary Schooling		-.252
		(.047)
Baseline Controls	Yes	Yes
R-Squared	.173	.178
Number of Observations	28,221	28,221

B. Health Expenditures

	Post-Period Health Insurance Status	
Positive Pre-Period Health Expenditures	-.043	-.062
	(.013)	(.018)
Treatment	.345	.336
	(.035)	(.038)
Treatment × Positive Pre-Period Health Expenditures		.036
		(.029)
Baseline Controls	Yes	Yes
R-Squared	.173	.173
Number of Observations	28,394	28,394

Notes: Entries are coefficients and standard errors from estimating equation (1) by ordinary least squares. The omitted category in the upper panel is 'primary school or less'. Heteroskedasticity robust standard errors are clustered on the level of randomization and reported in parentheses. In addition to the variables shown in the table, indicator variables for missing values on each covariate are also included in the regressions. See the Data Appendix for the precise definition and source of each variable.