ECONSTOR
Make Your Publications Visible.

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Neelsen, Sven; O'Donnell, Owen

Working Paper

Progressive Universalism? The Impact of Targeted Coverage on Healthcare Access and Expenditures in Peru

Tinbergen Institute Discussion Paper, No. 16-019/V

Provided in Cooperation with:
Tinbergen Institute, Amsterdam and Rotterdam

Suggested Citation: Neelsen, Sven; O'Donnell, Owen (2016) : Progressive Universalism? The Impact of Targeted Coverage on Healthcare Access and Expenditures in Peru, Tinbergen Institute Discussion Paper, No. 16-019/V, Tinbergen Institute, Amsterdam and Rotterdam

This Version is available at:
https://hdl.handle.net/10419/130499

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

TI 2016-019/V
Tinbergen Institute Discussion Paper

Progressive Universalism? The Impact of Targeted Coverage on Healthcare Access and Expenditures in Peru

Sven Neelsen[a]
Owen O'Donnell[b]

[a] *Institute of Health Policy & Management, Erasmus University Rotterdam, the Netherlands;*
[b] *Erasmus School of Economics, Erasmus University Rotterdam, Tinbergen Institute, the Netherlands; University of Lausanne, Switzerland; University of Macedonia, Greece.*

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

More TI discussion papers can be downloaded at http://www.tinbergen.nl

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 1600

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Progressive universalism? The impact of targeted coverage on healthcare access and expenditures in Peru

Sven Neelsen[a*] and Owen O'Donnell[b,c,d,e**]

[a] *Institute of Health Policy & Management, Erasmus University Rotterdam*

[b] *Erasmus School of Economics, Erasmus University Rotterdam*

[c] *Tinbergen Institute, the Netherlands*

[d] *Faculty of Business and Economics, University of Lausanne*

[e] *School of Economics and Regional Studies, University of Macedonia, Greece*

March 2016

Abstract

Like other countries seeking a progressive path to universalism, Peru has attempted to reduce inequalities in access to healthcare by granting the poor entitlement to tax-financed basic care without charge. We identify the impact of this policy by comparing the target population's change in healthcare utilization with that of poor adults already covered through employment-based insurance. There are positive effects on receipt of ambulatory care and medication that are largest among the elderly and the poorest. The probability of getting formal healthcare when sick is increased by almost two fifths, while the likelihood of being unable to afford treatment is reduced by more than a quarter. Consistent with the shallow cover offered, there is no impact on use of inpatient care. Mean out-of-pocket (OOP) expenditure on healthcare is unaffected but spending is reduced by up to one quarter at some points of the distribution. Among healthcare users, medical spending is reduced across much of the distribution and in relative terms falls most at lower quantiles, which is consistent with limited nominal and effective coverage of expensive treatments.

JEL Classification: H42, H51, I18

Keywords: Health insurance, health financing, healthcare, Universal Coverage, Peru

Acknowledgements: Sven Neelsen received funding from the EU FP7 Marie Curie Intra European Fellowship HEALTHEVENT (PIEF-GA-2012-331815). He thanks the Economics Department of the University of Piura and Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) for facilitating a research stay in Lima. We are indebted to Zósimo Juan Pichihua Serna and Luis Alexander Orrego Ferreyros for their insights into the Peruvian health system, and to Noelia Bernal Lobato, Miguel Ángel Carpio, José Luis Chumbe Casas and Pilar García Gómez for valuable discussions. We are grateful for comments received at seminars at Erasmus, Harvard, Lancaster, Lugano and Piura, as well as the XXXIV Jornadas Economía de la Salud (Pamplona).

* Corresponding author: Institute of Health Policy and Management, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, the Netherlands, E-mail address: neelsen@bmg.eur.nl

** odonnell@ese.eur.nl

1. Introduction

Provision of basic healthcare for the poor is proposed as a stepping stone on a bottom-up path to universal health coverage (Cotlear et al 2015). It aims to direct resources of severely constrained public health systems to the priority of improved access for the poor. Yet this form of *progressive universalism* (Jamison et al. 2013; Gwatkin and Ergo 2011) can be thrown off track by inaccurate targeting of the poor, ineffective fee waivers, inadequate budgets that leave underfunded and understaffed facilities with little capacity to respond to increased demand, sparse provision in rural areas and a narrow benefit package that does nothing to reduce the risk of immiserizing payments for hospital care. Free public healthcare for the poor cannot be presumed to deliver on the promise of more equitable access and improved protection from medical expenditure risks, and there is relatively little evidence on the extent to which it does.

This paper evaluates an attempt to reduce inequality in access by providing poor Peruvian adults with fully subsidized cover for basic healthcare at public facilities. Potential brakes on the effectiveness of the programme included reliance on a proxy means test to identify the poor, tight supply-side constraints due to manpower shortages and inadequate provision in rural areas, and weak incentives for facilities to meet demand arising from the removal of user fees that were previously an important source of revenue.

We identify the impact of the reform by comparing the target population's change in healthcare utilization and out-of-pocket expenditure with those of poor adults who were already covered through employment-based insurance. We find positive effects on receipt of ambulatory care, medication and diagnostic tests that are large relative to baseline rates. The estimates are generally robust to identification by either linear or nonlinear difference-in-differences and to alternative methods of conditioning on covariates. There is no evidence that the treatment and comparison groups were following different trends prior to the reform. The impacts increase over time, suggesting that the reform took a couple of years to become fully effective. They are larger in the older and the poorest populations that are likely to have had the greatest unmet need for care prior to being covered.

There is no significant impact on receipt of inpatient treatment. This is consistent with the exclusion of much tertiary care from the benefit package and with ceilings placed on coverage, but it may also be attributable to the weak incentives for providers to deliver high cost procedures at a low rate of reimbursement. Supply constraints appear to have left effective coverage short of nominal coverage for higher level treatments.

The increased use of ambulatory care appears to be due to effective implementation of the mandated exemption of the poor from fees at public facilities. The probability that a poor person

is treated by a formal healthcare provider when sick is increased by almost two fifths, while the probability that financial circumstances prevent the infirm from seeking formal treatment is reduced by more than a quarter. Despite this, there is no impact on mean out-of-pocket (OOP) medical expenditure. This is likely due to the offsetting effect of increased utilization of treatments that are only partially covered. There are significant reductions of 17-26% around the third quartile of the distribution of OOP expenditure. Among users of healthcare, spending is reduced at most points in the distribution and, consistent with limited cover of high cost tertiary treatment, the reduction is larger in relative terms at lower quantiles.

Bernal, Carpio, and Klein (2015) evaluate public healthcare for the poor in Lima four years after the programme was introduced nationwide, and after deepening of the benefit package and revision of the means test. Using variation in entitlement around the poverty threshold, they find positive effects on use of ambulatory care and medicines that are larger than our estimates. Unlike the present paper, they detect a significant positive impact of the more generous coverage on inpatient treatment. Also contrary to our findings, they estimate a positive effect on household OOP expenditure and attribute this to improved access to healthcare that increases awareness of medical needs and, consequently, willingness to pay for medicines and treatments that are not fully covered.

This paper complements Bernal, Carpio, and Klein (2015) in a number of respects. The latter, like evaluations of public health insurance for the poor in Georgia (Bauhoff, Hotchkiss, and Smith 2011) and Colombia (Miller, Pinto, and Vera-Hernández 2013), identifies the effect of cover on individuals at the margin of poverty. This is relevant to the extension of coverage to near poor populations, which is on the policy agenda of many upper middle-income countries. The evidence we present on the impact of covering the poor adult population is of greater relevance to countries that are contemplating taking this step and are looking to learn from those that have.

By restricting attention to Lima, Bernal et al do not provide evidence on the impact of the programme in parts of the country where need is substantially greater and medical facilities and manpower are considerably more sparse than they are in the capital city. Mexico's public health insurance for the poor appears to have reduced OOP medical expenditures only in urban centres and in the rural areas with more health facilities (Grogger et al. 2015). Our point estimates suggest smaller impacts on utilization in rural areas, although the differences are not significant.

Bernal et al identify the eligible through a proxy mean test that was administered by an independent agency. In the period we study, the means test was conducted on presentation for treatment at a health facility. While this is more prone to manipulation, it is typical of the procedure used to establish entitlement to fee exemptions in many lower income settings. We estimate intention-to-

treat effects on the policy's target population – poor adults identified according to the official measure of poverty. These estimates capture both the efficiency of the targeting achieved by the means test and the impact on the poor who are granted entitlement.

There is surprisingly little credible evidence on the impact of fully subsidized public healthcare for the poor.[1] An Indonesian scheme that provides the poor with access to basic healthcare at public facilities without charge is found, like the similar Peruvian programme we evaluate, to increase use of ambulatory care (Sparrow, Suryahadi, and Widyanti 2013). Less consistent with our results, but similar to Bernal, Carpio, and Klein (2015), this study finds some evidence of a positive impact on OOP spending, particularly in urban areas and among richer households that manage to enroll despite the targeting. Health insurance for the poor in Colombia, which differs markedly from the Peruvian and Indonesian variants, is found to raise preventive care and consultations when sick, but not hospitalization rates (Miller, Pinto, and Vera-Hernández 2013). The level and the variance of OOP spending are reduced, consistent with insurance offering substantial protection against medical expenditure risk. A similar managed competition insurance programme in Georgia that has no copayments for medical care but excludes drugs from the benefit package also appears to reduce mean OOP payments and the risk of heavy spending on inpatient treatment but with no impact on utilization (Bauhoff, Hotchkiss, and Smith 2011). Mexico's Seguro Popular, which in principle offers fully subsidized health insurance only to the poor but in practice covers all who enrol without charge, has also been found to reduce OOP spending without impacting on healthcare utilization (King et al. 2009). The same conclusion is reached with respect to the impact of health insurance for the poor in Vietnam (Wagstaff 2010).

Despite the controversy generated by user fees, there is even less sound evidence on their effects than there is for targeted health insurance (Lagarde and Palmer 2008).[2] We know of no study that convincingly identifies the impact of individual, as opposed to geographic, targeted removal of user fees.[3] Before-and-after studies generally show increased utilization of outpatient but not inpatient care after universal removal of user fees (ibid, Table 7). Comparison between rural districts of Zambia in which fees were removed and (peri-) urban areas in which they were not (until later) reveals no impact on utilization but the near elimination of OOP medical expenditures (Lépine, Lagarde, and Le Nestour 2015).

[1] We consider only studies with a credible strategy to identify the impact of a nationwide programme offering fully subsidized coverage to poor adults identified through an explicit targeting mechanism, such a proxy mean-test.

[2] Prior to user fees for public healthcare being levied in Peru, Gertler, Locay, and Sanderson (1987) argued that they would result in a substantial welfare loss due to a shift in the health financing burden from richer taxpayers to poorer users. They argued in favour of the policy we evaluate – user fees contingent on ability to pay.

[3] Flores et al. (2013) find that fee waivers made effective by compensating facilities for lost revenue reduce OOP payments in Cambodia. A randomized experiment that removed user fees for households with young children in one district of Ghana found health gains – raised haemoglobin of anaemic children (Powell-Jackson et al. 2014).

Overall, evidence on the impact of subsidized healthcare for the poor in low- and middle-income countries is scant and contradictory. We add evidence on Peru's initial targeted coverage programme that possessed many characteristics typical of attempts to follow a bottom-up path to universal coverage (Cotlear et al 2015). Rather than estimate, as previous studies, the effect on those obtaining coverage by passing a proxy means test, we estimate the impact over the target poor population. This is the effect most relevant to a government aiming to improve access to healthcare and provide protection against medical expenses among the poor.

2. Health care and insurance in Peru

2.1. Health insurance before the 2007 reform

At the time of the coverage extension, Peru was an upper-middle income country spending only 4.5% of GDP on healthcare (World Health Organization 2016). Just over half (54%) of this expenditure was publicly financed. The greatest part of the remainder (40%) was paid for out-of-pocket (ibid). Over three fifths of the 28 million population had no health insurance cover (INEI 2006).[4] They paid out-of-pocket for public and private healthcare, or went without treatment. One quarter of individuals (and one third of the poor) who reported not to have received formal medical treatment when sick claimed the reason was financial (INEI 2006).

From 2002, poor children (<18 years) and poor pregnant women were exempted from paying user fees for basic healthcare at National Health Service (NHS) facilities.[5] Providers were reimbursed through the tax-financed Seguro Integral de Salud (SIS), which covered 16% of the population in 2006. This programme also covered basic emergency care for poor uninsured adults with life-threatening or potentially permanently damaging conditions.

The programme is not an autonomous fund that purchases healthcare on behalf of its beneficiaries. Rather, it is one of the budgetary channels through which the NHS is financed (World Bank 2011). The budget is set annually by the Ministry of Finance on an ad hoc basis rather than in direct relation to the population covered. In 2006, the budget allowed spending of just $18 per beneficiary. During the period studied, facilities were reimbursed according to tariffs set to cover the costs of medicines and consumables, plus a markup. Salary and capital costs were not reimbursed. They had to be covered from the general NHS budget. Claims usually exceeded the

[4] Estimated from the nationally representative household survey, ENAHO.

[5] Also exempted were citizens active in social grass-root organizations. From 2006, victims of human rights violations, certain ethnic minorities, shoe shiners, the homeless and community health workers also gained coverage. Even so, poor children and poor pregnant women constituted 96% of those enrolled and accounted for 97% of programme spending in 2006 (SIS 2007).

programme's regional budget resulting in delay or outright refusal of reimbursement (World Bank 2011; Francke 2013).

There was and remains a gross shortage of medical manpower. In 2006, there were only 1.24 health workers in public facilities per 1,000 population (Ministerio de Salud 2015). This is well below the recommended threshold of 2.3 for effective delivery of primary care (World Health Organization 2013). The average masks stark inequality to the disadvantage of rural areas – the medical worker-population rate ranges from 0.7 to 3.7 across 25 regions.

Before and after the 2007 reform, formal sector employees and their dependents were covered by the *El Seguro Social en Salud* (EsSalud). Coverage through this scheme is retained after retirement and for up to one year of unemployment. Finance is by a 9% payroll tax. Independent workers can enroll voluntarily but very few do.[6] The programme covered 18% of the population in 2006. The benefit package comprehensively covers ambulatory and inpatient care, including high-cost treatments, as well as medicines. There is no cost-sharing. Care is accessed through the scheme's own provider network that does not overlap with the NHS facilities accessed through the programme for the poor. Spending per beneficiary was $205 in 2006, one third above the national average.[7] Still, resources were strained and long waiting times resulted in a non-trivial share of beneficiaries opting to pay OOP for direct access to private health services (INEI 2006).

2.2. The 2007 reform

Legislation extending coverage through the Seguro Integral de Salud (SIS) programme to the entire poor population was passed in June 2005 but it was not put into effect until March 2007. This gave six million Peruvian adults – 21% of the population – entitlement to basic healthcare at public NHS facilities without charge.[8] Previously, these individuals were required to pay user fees ranging from $2 for an outpatient consultation to $53 for a hospitalization with major surgery, not including payments for medicines and medical supplies.

Simultaneous to the extension of population coverage, the SIS benefit package was made more comprehensive. In principle, it fully covered treatments for conditions estimated to account for around a quarter of the burden of disease (Francke 2013). Treatment of other conditions was supposedly covered up to a ceiling of $2875 per episode, and a high-cost disease fund provided

[6] In 2006, 1.8% of EsSalud beneficiaries were covered through one of its subsidized voluntary options (EsSalud 2008).
[7] Calculated from total expenditure given in 2006 annual report (EsSalud 2007) divided by the beneficiary count (5.1 million) estimated from 2006 ENAHO survey data. Using the official 2006 beneficiary count of 6.4 million gives $162 per capita.
[8] The reform also established a semi-contributory option for the near-poor. Despite heavily subsidized premiums, uptake has been minimal. Not even 1% of SIS beneficiaries were covered through this option in 2010 (SIS 2011).

additional coverage for treatments costing up to $5750.[9] There were no copayments for treatments in the benefit package. However, lack of budgeting for the projected costs of delivering the benefit package, the continued exclusion of salary costs from the reimbursement rates, poor access to facilities in rural areas, manpower shortages, ineffective medicines procurement and widespread lack of inputs necessary for surgery mean that the service coverage actually available is likely to have fallen well short of that promised (Francke 2013; Vermeersch, Medici, and Narvaez 2014; Bernal, Carpio, and Klein 2015). Effective coverage of inpatient treatment was likely to have been particularly limited.

Two strategies were used to enroll the poor population. In the poorest 880 (/1,832) districts, every uninsured individual could enroll when accessing a NHS facility. In the rest of the country, as had been the case for children and pregnant women prior to the reform, the poor were identified by a proxy means test. This was administered at an NHS facility on the basis of information on asset ownership, housing conditions, demographics, education and occupation provided by the applicant. Poverty status was determined by an algorithm that was kept secret to reduce the risk of gaming. Enrolment was immediate if there was a computer at the facility and otherwise took up to eight weeks. It was free of charge and remained valid for three years.

Means testing at the facility is potentially vulnerable to both adverse selection by non-poor claimants in bad health and strategic rejection of poor applicants by providers seeking to protect user fee revenue given modest and unreliable reimbursement from the insurance programme. Facilities are more certain of user fee income and have greater autonomy in spending it (Francke 2013). Estimates of the programme's target efficiency vary. Leakage to the non-poor and better-off is estimated to be 12-14% (Arróspide, Rozas, and Valderrama 2009; Vermeersch, Medici, and Narvaez 2014). The gap in coverage among the poor is estimated in the range of 16-24% (Arróspide, Rozas, and Valderrama 2009; Vermeersch, Medici, and Narvaez 2014). These inaccuracies reflect not only strategic behaviour on the part of both patients and providers but also the inevitable discrepancy between a proxy means test and a more complete measure of poverty based on consumption data (Vermeersch, Medici, and Narvaez 2014).

Figure 1 shows coverage rates based on health insurance status reported in the nationally representative ENAHO household survey. From 2007, enrolment in the programme for the poor rises and there is a corresponding fall in the estimated fraction of the population without cover

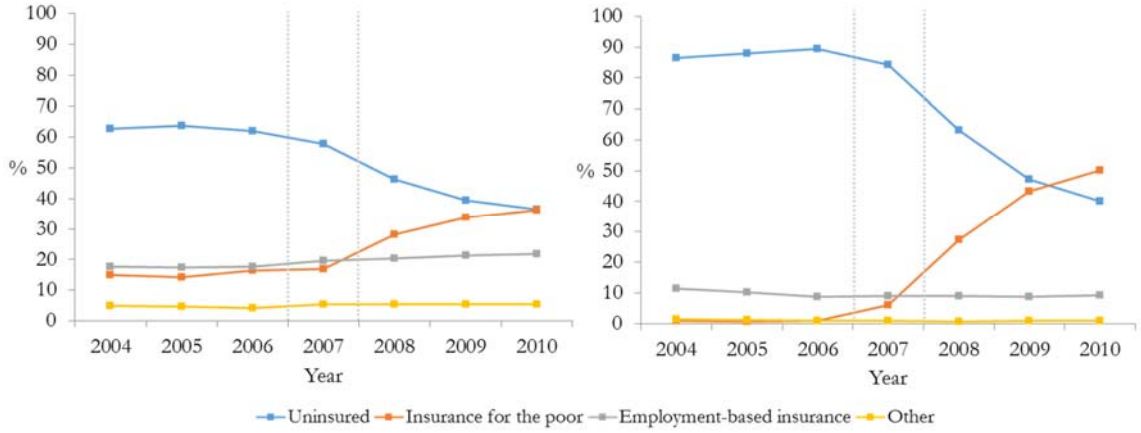
[9] Conditions not in the core benefit package but covered up to the ceiling included asthma, benign breast tumours, diseases of the urinary tract, gallstones, skin infections, eye and eyelid infections, epilepsy, back pains, early detection of diabetes mellitus and dyslipidaemia, and neoplasms. The coverage ceiling was $1725 for the treatment of neoplasms. The benefit package was more comprehensive for those aged below 20 years.

from 62% in 2006 to 46% in 2008. Over the next two years, the share of the sample that reports being uninsured falls by an additional ten percentage points. In the target population of poor adults, which we identify using the official poverty line, the fraction uninsured is estimated to fall from 90% in 2006 to 40% by 2010. In that year, one half report being enrolled in the health insurance for the poor programme. Only one percent had been enrolled in 2006. Most likely, the gradual increase in reported coverage during the post-reform period is because potential beneficiaries postponed enrolment until treatment was needed, given cover took effect immediately. It may also be because many were unaware of their acquired entitlement due to limited promotion of the coverage extension (Arróspide, Rozas, and Valderrama 2009), or had been denied cover by facilities operating strategically.

Figure 1: **Health insurance coverage rates**

a) Full population

b) Target population (poor adults)



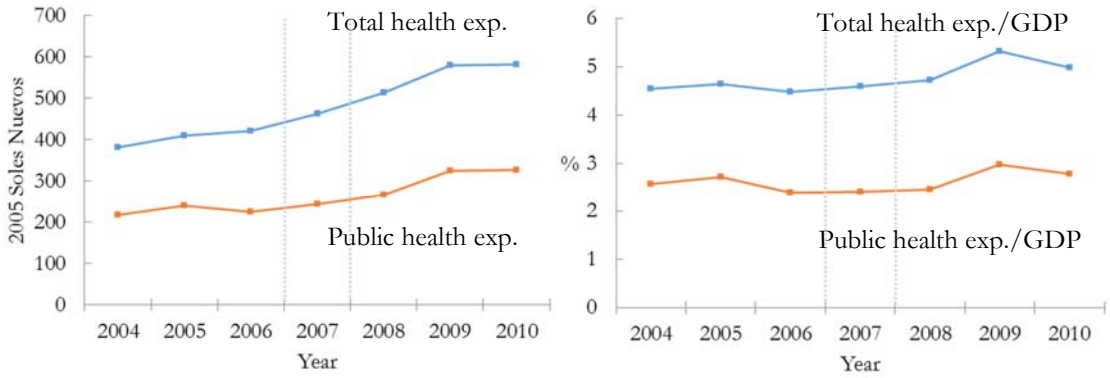
Notes: National household survey (ENAHO) estimates of percentage reporting each type of health insurance cover. Population weights applied. 'Insurance for the poor' is reported enrolment in the SIS programme. 'Employment-based insurance' is reported enrolment in the EsSalud scheme. 'Other' includes schemes for the military and police, plus private insurance. The target population consists of adults (> 19 years) with household per capita consumption below the official poverty line.

The coverage extension was accompanied by increased spending on healthcare (Figure 2). Due to rapid economic growth, health expenditure as a share of GDP held stable until 2008 before increasing in 2009 due to the sharp increase in real health expenditure and the stalling of the economy in the wake of the financial crisis. Real public health expenditure per capita increased by 18.7% between 2006 and 2008, compared with an increase of only 3.7% between 2004 and 2006 (World Health Organization 2016). By 2009, public expenditure was 44% above what it had been in 2006 and the budget of the SIS programme had risen by 47%. But this was not nearly proportionate to the expansion of population coverage, with the result that real programme

spending per person covered decreased by 31%.[10] The gap between nominal and effective coverage most likely widened. Implicit rationing was necessary to keep spending within the programme budget that was not set on the basis of a costed benefit plan and size of the covered population (Francke 2013).

Figure 2: Total and public health expenditure

a) Soles Nuevos per capita (2005 prices) b) Percentage of GDP



Source: WHO Global Health Expenditure Database (World Health Organization 2016).

3. Data

3.1. Sample

We use four waves of the nationally representative ENAHO cross-section survey conducted by Peru's National Institute of Statistics in the period 2005-2009. A consistent sampling design and questionnaire is used to survey around 20,000 households containing about 90,000 individuals annually. Interviews take places throughout the year.

The individual is the unit of analysis. While the reform increased the depth of coverage for population groups – principally, poor children and poor pregnant women – previously covered by the insurance for the poor programme, we restrict attention to the impact on those not previously covered – poor adults. We identify this target group as follows. First, we drop from the sample individuals who report private insurance cover, those below the age of 20 and women reporting to have used antenatal care or delivery assistance over the past 12 months.[11] Second, we restrict the sample to individuals identified as poor. Poverty status is assessed not on the basis of the proxy

[10] Own calculations based on SIS annual reports, www.sis.gob.pe/Portal/transparencia/organizacion.html.
[11] We also drop those with cover through the military and police forces. We exclude individuals aged 18 and 19 since the benefits they became entitled to in 2007 differed from those aged 20+. Our sample will include some pregnant women who did not report using antenatal care or birth attendance in the past 12 months although they were entitled to access these services even before the reform.

means test used at health facilities to assess eligibility but using the official definition of household per capita consumption below the regional poverty line.[12] This identifies the target poor population that would be covered by the programme if it were feasible to assess entitlement using the most comprehensive measure of poverty. Our goal is to estimate the intention-to-treat effect on this target population. The less efficient was the means test in targeting the poor and the less awareness there was of programme entitlement, the greater will be the divergence between our analytical sample and a random sample of those awarded coverage. This will not result in a biased estimate of the effect in the target population. Rather, it will correctly estimate an effect that is reduced by deficiencies in targeting the poor and alerting them of their entitlement.

The third restriction we impose on the sample is to drop observations from the 880 poorest districts where the automatic enrollment procedure applied. These districts benefited from the national poverty alleviation strategy (CRECER) consisting of numerous nutrition and social programmes implemented near simultaneously to the extension of health insurance coverage from 2007 (Borra 2010). While the anti-poverty programme focused on (malnutrition of) children and the healthcare coverage reform targeted poor adults, disentangling the effects of the two policies is difficult. We demonstrate robustness of our estimates to including these districts.

3.2. Treatment & comparison groups

Within the selected sample of poor adults, we assign those who report being covered by employment-based social health insurance (EsSalud) to the comparison group and all other individuals to the treatment group. The comparison group consists of low-paid salaried employees, poor retired and unemployed former salaried employees, certain categories of workers, and adult dependents of these groups.[13] Nominal and effective coverage for beneficiaries of the employment-based programme were stable over the 2005-9 period: the benefit package was unchanged and real expenditure per enrollee increased little from $198 in 2005 to $204 in 2009 (EsSalud 2006; EsSalud 2010).

The treatment group consists of individuals who report being covered by the health insurance for the poor programme or having no insurance. As is clear from Figure 1b, a substantial fraction of

[12] We diverge from the official measure by comparing household consumption net, not gross, of health payments with the poverty line. This avoids classifying individuals as non-poor because of high OOP spending financed from savings, credit, etc.. This does not rule out selection on the outcome if health spending is partly financed by cutting back on other consumption. There is some evidence that households can smooth consumption over temporary illness and medical treatments in Peru (Diaz and Valdivia 2011). Nevertheless, we check the robustness of our estimates to using total consumption gross of health payments to identify the poor.

[13] The comparison group may include some rural workers, domestic servants, fishermen and certain other independent workers who enrolled voluntarily in the social insurance programme at a reduced contribution rate. Such workers constituted only one percent of total enrolment in 2009 (EsSalud 2010). Our estimates are robust to excluding households with an independent worker from the comparison group.

the target group continues to report being uninsured after the reform. Besides being consistent with the identification of intention-to-treat effects, inclusion of these individuals in the treatment group avoids bias that would arise if enrollees were compared with non-enrollees given that enrolment can take place at the time of healthcare utilization.

Data from the 2005 and 2006 surveys are used to provide information on the pre-reform period and those from 2008 and 2009 document the post-reform period. Over the four cross-sections, we have 47,997 observations in the treatment group and 5,716 in the comparison group.

The left panel of Table 1 shows means of covariates in the pre-reform period for the treatment and comparison groups. This reveals substantial and statistically significant differences in sociodemographic characteristics that reflect the attachment of only one group to the formal economy. Among other things, the comparison group is older and has a larger share of individuals not working due to the inclusion of retired social insurance beneficiaries. It also has a larger share of married individuals because coverage extends to spouses. Because social insurance is tied to formal sector employment, the comparison group also has higher shares of employment in services and non-manual occupations, it is better educated, more likely to reside in urban areas and, on average, has a higher level of consumption (below the poverty threshold). Moreover, the comparison group enjoys better living conditions, with higher rates of access to improved water sources, sanitation and modern cooking fuels (see online Appendix Table A1).

The right-hand panel of Table 1 shows the pre-post reform change in the mean of each covariate for each group. There are fewer differences in the differences than there are in the levels. The treatment group somewhat catches up in mean age but the difference in the gender balance widens. There are also some differences in changes in education and service sector employment, and consumption growth is faster in the treatment group than it is in the comparison group.

Controlling for these time varying covariates will help reduce the risk of compositional bias in our difference-in-differences estimates. But given the baseline differences in observables, it bears emphasis that the estimates will only be informative of the impact of the coverage extension if the comparison group, conditional on covariates, is indicative of the outcomes that the treatment group would have experienced in the absence of the reform. To increase the plausibility of this identification condition, we demonstrate robustness of the estimates to both weighting the comparison observations to make them more similar to the treated and restricting the treatment group to observations that are more similar to the controls.

Table 1: Covariate means and changes in means by treatment status

	Pre-reform (2005/6)			Change (2008/9-2005/6)		
	Treatment (T)	Comparison (C)	T=C p-value	Treatment (△T)	Comparison (△C)	△T =△C p-value
Demographics						
Age in years	41.04	47.06	0.000	1.674***	0.522	0.044
Male	0.503	0.525	0.017	-0.007	0.038***	0.000
Married or legal partnership	0.651	0.748	0.000	0.009**	0.013	0.783
# of household members	5.619	5.927	0.000	-0.157***	-0.211***	0.550
Highest level of education						
None	0.105	0.062	0.000	0.018***	0.013*	0.573
Incomplete primary	0.227	0.176	0.000	0.018***	-0.004	0.059
Complete primary	0.182	0.140	0.000	-0.003	-0.003	0.988
Incomplete secondary	0.161	0.114	0.000	-0.007**	0.015	0.027
Complete secondary	0.220	0.270	0.000	-0.015***	0.001	0.230
Any tertiary	0.104	0.238	0.000	-0.010***	-0.021*	0.366
Employment sector						
Not working	0.225	0.315	0.000	-0.026***	-0.046***	0.131
Agriculture/fishing	0.358	0.133	0.000	0.023***	0.007	0.178
Mining/manufacturing/construction	0.103	0.139	0.000	0.015***	0.010	0.622
Services	0.314	0.414	0.000	-0.012***	0.029**	0.006
Occupation						
Professional	0.021	0.116	0.000	-0.003**	-0.016*	0.148
Clerical support, service, or sales worker	0.093	0.148	0.000	-0.000	0.010	0.353
Skilled agricultural/forestry/fishery worker	0.173	0.058	0.000	0.019***	0.012*	0.343
Craft and related trades worker	0.063	0.081	0.000	0.006**	0.002	0.687
Plant/machine operator, assembler	0.056	0.050	0.139	0.012***	0.015**	0.623
Elementary	0.370	0.231	0.000	-0.007	0.023*	0.023
Log monthly non-health consumption p.c.	4.900	5.178	0.000	0.035**	0.013	0.060
Urban (>400 homes)	0.619	0.858	0.000	-0.013***	0.004	0.207
Number of observations	*28,073*	*3,470*				

Notes: Sample of poor individuals aged 20+ who do not report use of maternity care in past 12 months and reside outside the 880 poorest districts. Treatment group (T) consists of those reporting cover through health insurance for the poor or no insurance. Comparison group (C) includes those reporting employment-based health insurance. ***, **, * indicate that that change over time is significant at 1%, 5% and 10% respectively. Means for other covariates included in the estimation models given in Appendix Table A1.

4. Impact on healthcare utilization

4.1 Outcomes and simple difference-in-differences

The reform was intended to improve the poor's access to healthcare. We estimate the impact on the following indicators of utilization: i) any ambulatory consultation in the last four weeks; ii) any use of medicines (prescribed or not) in the last four weeks; iii) any diagnostic test or examination in the last four weeks; iv) any visit to dentist or ophthalmologist in the last 3 months; and, v) any hospitalization or surgical operation in the last year.[14] Treatments within each of these broadly defined health services are included in the benefit package of the health insurance for the poor programme, although not all treatments within each category are covered. For example, an ambulatory consultation for an acute infection would be covered but treatment for a sequela of that condition would not, unless it was immediately life-threatening. Only medication issued in

[14] The information is provided by the survey respondent for all household members.

NHS facilities is covered. Nominal coverage of dental and ophthalmic care is basic and effective coverage likely even more limited due to a shortage of supply. The programme only covers treatment in NHS facilities, which is not distinguished from treatment in other facilities in the data. This is not a limitation since the aim is to establish whether the reform succeeding in increasing access. If it merely resulted in the substitution of public for private treatment, then no effect will be found.

Table 2 shows the mean of each utilization indicator in the pre- and post-reform periods for the treatment and the comparison group, as well as the simple difference-in-differences across the groups. Prior to the reform, all indicators show substantially lower utilization in the treatment group that lacked insurance compared with the comparison group that did not. Less than 15% of the treatment group had at least one ambulatory consultation in the past four weeks – approximately half the rate of the comparison group. The rate of use of dental or ophthalmic care of the treatment group was also around half of that of the comparison group, while the treatment group's likelihood of getting a diagnostic test was around a third of that of the comparison group. Only 3.6% of the treatment group had an inpatient stay in the last year compared with 6.7% of the comparison group. The smallest relative difference prior to the reform was in utilization of medicines. But even in this case the treatment group was ten percentage points less likely to receive medicines in the last month. All of the baseline differences in the means across groups are statistically significant at 1%.

The third column of Table 2 reveals significant increases in utilization of all types of healthcare except for dental and ophthalmic care in the group to which coverage was extended. For the comparison group, in which coverage was unchanged, there was a significant increase only in ambulatory care and in inpatient care. For ambulatory care, the increase in use by the comparison group is less than half of that by the treatment group. As a result, the simple difference-in-differences (DID) indicates a significant five percentage point rise in the probability of having a consultation, which is attributable to the reform under the assumption that the trends in the utilization rates would have been equal across the groups in the absence of the reform. This is more than a third of the baseline rate of utilization of the treatment group. The simple DID estimate of a significant increase of 4.6 percentage points in the probability of receiving medication is 14% of the baseline rate. For diagnostic tests and dental/ophthalmic care, the point DID estimates are both positive but not significant. The probability of a hospitalization increased by almost twice as much in the comparison group as in the treatment group and so the unconditional DID is negative but not remotely significant.

Overall, comparison of the changes in means suggests that the reform may have raised utilization of ambulatory care and medicines among the poor adult population previously without cover but that it did not effectively increase access to inpatient treatment.

Table 2: Healthcare utilization rates

	Pre-reform (2005/6)	Post-reform (2008/9)	Difference (post-pre)	Difference-in-differences	*Sample size (by group)*
Ambulatory consultation in last 4 weeks					
Treatment	0.145	0.230	0.085***	0.050***	*47,997*
	(0.002)	(0.003)	(0.004)	(0.014)	
Comparison	0.288	0.324	0.036***		*5,716*
	(0.008)	(0.010)	(0.013)		
Medicines in last 4 weeks					
Treatment	0.332	0.389	0.058***	0.046***	*47,997*
	(0.003)	(0.004)	(0.005)	(0.015)	
Comparison	0.435	0.447	0.012		*5,716*
	(0.009)	(0.011)	(0.014)		
Diagnostic tests or examinations in last 4 weeks					
Treatment	0.025	0.035	0.011***	0.008	*47,997*
	(0.001)	(0.001)	(0.002)	(0.008)	
Comparison	0.078	0.081	0.003		*5,716*
	(0.005)	(0.006)	(0.007)		
Visit to dentist or ophthalmologist in last 3 months					
Treatment	0.063	0.065	0.002	0.001	*47,997*
	(0.002)	(0.002)	(0.002)	(0.009)	
Comparison	0.124	0.125	0.001		*5,716*
	(0.006)	(0.007)	(0.009)		
Inpatient stay or surgery in last 12 months					
Treatment	0.036	0.044	0.008***	-0.006	*47,997*
	(0.001)	(0.001)	(0.002)	(0.008)	
Comparison	0.067	0.082	0.015**		*5,716*
	(0.004)	(0.006)	(0.007)		
Sample size (by period)	*31,543*	*22,170*			*53,713*

Notes: Sample as well as treatment and comparison groups defined as in notes to Table 1. Robust standard errors adjusted for clustering at the household level in parentheses (Rogers 1994). All differences between the treatment and comparison groups in both the pre- and post-reform periods are significant at the 1% or 5% level. ***, **,* indicate significant post-pre difference and difference-in-difference at 1%, 5% and 10% respectively.

While the pre-reform differences in utilization rates between the treatment and comparison groups do not invalidate the difference-in-differences identification strategy, the magnitude of the differences, together with those in covariates documented in Table 1, may give cause to question the plausibility of the common trends assumption. We test whether the outcomes were following common trends in the treatment and comparison groups prior to the reform. We use data from the 2004, 2005 and 2006 ENAHO surveys, select the sample and define treatment and comparison groups as above and test whether the change in utilization rate in 2005 and in 2006 was equal between these groups. There is no differential trend in either ambulatory care or medication (see online Appendix Table A2, columns indicating no control for covariates). This lends credibility to

interpretation of the simple DID estimate as indicative of a positive impact of the reform on access to each of these types of care. The only period effects that differ significantly across the groups are for dental/ophthalmic care (2005) and inpatient care (2006). But for neither of these types of care is there a significantly positive DID estimate that we risk incorrectly attributing to the reform.

4.2 Conditional difference-in-differences

Controlling for time varying covariates can help reduce the risk of compositional bias and increase the precision of the estimates. Given all the outcomes are binary and some have low prevalence rates, we estimate the conditional probabilities using logit models specified as follows,

$$P\left(y_{it}=1\right)=\Lambda\left(\theta Treat_i+\tau_t Year_t+\delta Treat_i\times Post_t+\mathbf{X}_{it}\boldsymbol{\beta}\right) \tag{1}$$

where $y_{it}=1$ indicates that individual *i* in year *t* makes use of one of the health services, Λ is the logistic function, $Treat_i=1$ indicates membership of the treatment group, $Year_t=1$ indicates the observation is from year *t*, $Post_t=1$ if the observation is from year 2008 or 2009 and $\mathbf{X}_{it}$ is a vector of covariates. The latter includes all the variables listed in Table 1 and Appendix Table A1. Demographics are represented by 12 age-sex group indicators and urbanization of the location is captured by 8 indicators of housing density. In addition, we include indicators of 25 regions and of the month of the survey interview.

If, in the absence of the reform, all the year coefficients in the linear index would be common across the treatment and comparison groups, then the intention-to-treat effect is given by the partial effect of the *Treat*×*Post* indicator averaged over the (target) treatment group during the post-reform period (Puhani 2012),

$$ITT^y=\frac{1}{N_{TP}}\sum_i 1\left(Treat_i\times Post_t=1\right)\left[\Lambda\left(\hat{\theta}+\hat{\tau}_t Year_t+\hat{\delta}+\mathbf{X}_{it}\hat{\boldsymbol{\beta}}\right)-\Lambda\left(\hat{\theta}+\hat{\tau}_t Year_t+\mathbf{X}_{it}\hat{\boldsymbol{\beta}}\right)\right] \tag{2}$$

where 1() is the indicator function and N_{TP} is the number of observations in the treatment group in the post-reform period. Standard errors are computed by a delta method that is robust to heteroscedasticity and clustering within households, and takes account of the sampling variability of the covariates (Korn and Graubard 2011).

These estimated effects are presented in column (1) of Table 3. The estimated impact on ambulatory care is even larger than the simple DID estimate and suggests that the coverage extension increased the probability of any consultation within a four-week period by 5.9 percentage points, which is 41% of the pre-reform rate for the treated group of poor adults previously without insurance (Table 2).

It is estimated that the reform raised the probability of making use of medication within a four-week period by 3.8 percentage points, somewhat less than suggested by the simple DID estimate. The magnitude of the estimated effect on the probability of getting a diagnostic test, which is one third of the baseline rate, is not affected by conditioning on covariates but it becomes strongly significant. As with the simple DID estimate, there is no significant impact on the chance of visiting a dentist or ophthalmologist. The point estimate of the effect on the probability of a hospital admission remains negative but it falls in magnitude and remains not remotely significant.

These estimates suggest that the reform increased access to ambulatory care and medication, and also diagnostic testing, to an extent that is substantial relative to the rates at which poor adults were using these types of medical care prior to acquiring basic insurance cover. There was no impact on care that the programme covered only to a limited extent – inpatient, dental and ophthalmic treatment.

Table 3: Logit estimated of effects on healthcare utilization rates and robustness checks

Main estimate	Robustness to using		
	only 2006 (pre) & 2008 (post) data	official measure to select poor	all districts, incl. 880 poorest
(1)	(2)	(3)	(4)
Ambulatory consultation			
0.059***	0.044***	0.054***	0.053***
(0.009)	(0.014)	(0.011)	(0.009)
Medicines			
0.038***	0.018	0.039**	0.035***
(0.005)	(0.019)	(0.016)	(0.012)
Diagnostic test or examinations			
0.008***	0.010**	0.004	0.005*
(0.003)	(0.004)	(0.003)	(0.003)
Visit to dentist or ophthalmologist			
0.002	0.005	0.002	0.005
(0.006)	(0.007)	(0.006)	(0.004)
Any hospitalized or surgery			
-0.001	-0.012	-0.000	0.001
(0.005)	(0.009)	(0.005)	(0.004)
Sample size	*53,713*	*26,757*	*46,226*	*88,735*

Notes: Estimates of ITT effects defined in equation (2). In column (1), the sample is as defined in the notes to Table 1. In column (2), only data from 2006 (pre-reform) and 2008 (post-reform) are used. In column (3), the sample of poor adults is selected using household consumption gross (rather than net) of health payments to assess poverty status. In column (4), the sample is expanded to include the 880 poorest districts in which the CRECER poverty alleviation programme was implemented. All models include the covariates listed in Table 1 and A1, plus 12 age-sex group indicators, 25 region indicators, eight housing density indicators, year effects and month effects. Robust standard errors adjusted for clustering within households in parenthesis (Korn and Graubard 2011). ***, **,* indicate statistical signif. at 1, 5 and 10%.

The conditions required to identify the policy impact are that the logit function correctly specifies the conditional expectation of the outcome and that under the counterfactual of no reform the time coefficients in the linear index would be equal across the treatment and comparison groups (Puhani 2012). In the pre-reform period (2004-06), the latter restriction is not rejected for each of

the three types of care for which the conditional DID estimate is significant (see Appendix Table A2 columns indicating control for covariates). Only for dental/ophthalmic care is there a significantly different year coefficient, but the conditional DID estimate is not significant for this category in any case.[15]

4.3 Impact on unmet need for treatment

A positive impact of insurance on healthcare utilization may raise or lower welfare depending on whether it arises from an income or a pure price effect (Nyman 1999). In poor, credit-constrained populations, there is likely to be a strong access motive for insurance. Redistribution from the healthy to the sick can provide access to otherwise unaffordable but needed healthcare. Moral hazard – price induced utilization of low value healthcare – is of less concern. While it is very difficult to discriminate between moral hazard and income effects, the data do provide some opportunity to explore whether the extension of coverage reduced unmet need for treatment.

Table 4: Effects on reported illness and healthcare access conditional on illness

	All	*Observations reporting illness*	
	Report illness in last 4 weeks	Received formal treatment	Did not receive formal treatment due to finances
	(1)	(2)	(3)
	0.016	0.067***	-0.080***
	(0.014)	(0.011)	(0.030)
Baseline (treated)	0.555	0.174	0.301
Sample size	*53,713*	*31,276*	*31,276*

Notes: DID estimates of effects derived from logit models using equation (2). Sample (in column (1)) defined as in notes to Table 1. Covariates as in notes to Table 3. Robust standard errors corrected for within household clustering in parenthesis (Korn and Graubard 2011). ***, **,* indicate statistical significance at 1, 5 and 10%.

The first column of Table 4 shows the logit DID estimate of the impact of the reform on the probability of reporting illness in the previous four weeks.[16] It is not remotely significant and is small relative to the baseline incidence of illness. This suggests that there is unlikely to be much selection bias in the DID estimate of the impact on the probability of receiving formal treatment when sick, which is presented in the second column.[17] The estimate indicates a seven percentage point increase in the probability of treatment, which is almost two fifths of the baseline utilization

[15] An alternative way of testing common trends in the pre-reform period is to use the 2004 and 2006 data to obtain a logit DID estimate of the 'effect' of a placebo reform purportedly implemented in 2005. For each type of care, this estimate is small in magnitude and never close to being statistically significant (see online Appendix Table A3).

[16] Illness is defined as experiencing any of the following in the last four weeks: a) disease symptoms (cough, headache, fever, nausea), b) illness (flu, colitis, etc.), c) chronic disease relapse, d) accident.

[17] Selection corrected (heckprobit) estimates of effects on the probability of formal treatment and forgone treatment for financial reasons when sick are extremely close in magnitude and significance to those presented in Table 4. In the absence of plausible exclusion restrictions, identification is through nonlinearity.

rate.[18] The third column shows that the probability of not receiving formal treatment when sick because of lack of insurance or money is reduced by eight percentage points, or more than a quarter.[19] This is consistent with the insurance programme making needed medical care affordable to poor adults.

4.4 Robustness

The results presented in columns (2)-(4) of Table 3 demonstrate that the estimates are largely robust to alternative choices that could be made in selecting the sample. For the main analysis, we use a two-year window around the 2007 reform. Given the gradual increase in enrolment in health insurance for the poor observed in Figure 1, a two-year post-reform period increases the likelihood of capturing the full effect of the reform. Moreover, because the comparison group is rather small in comparison with the treatment group, a wider window makes the estimate less susceptible to statistical aberrations. On the other hand, a narrower window would reduce the risk of compositional bias.[20]

Column 2 of Table 3 shows estimates obtained using the 2006 and 2008 observations only. The point estimate of the effect on ambulatory consultations diminishes by about 25% in size relative to the main estimate but it remains highly significant and also large relative to the consultation rate prior to the reform.[21] The estimated impact on the utilization of medicines is less than half of that obtained using all four years of data and statistical significance is lost. This may reflect substitution of over-the-counter medicines with prescription drugs immediately after coverage of the latter was introduced with a positive net effect only emerging later. The positive, significant impact on diagnostic testing is somewhat larger with the narrower estimation period. There remains no significant effect on dental/ophthalmic care and on inpatient care, although for the latter there is a substantial rise in the magnitude of the negative point estimate.[22]

Poverty is officially defined as total household consumption below a threshold. We depart from this by identifying the poor as those with consumption net of health payments below the poverty

[18] Formal treatment is defined as receipt of care in response to illness or accident at any on nine types of public and private healthcare facilities as opposed to being treated at a pharmacy/drugstore, a traditional healer, home or not seeking care at all.
[19] Those reporting illness but not having received formal treatment are asked why they did not visit a health centre or facility. We identify those answering that they did not have insurance or money.
[20] The changes between 2006 and 2008 in the means of the covariates are more similar across the treatment and comparison groups than are the changes between 2005-06 (pooled) and 2008-09 (see Appendix Table A4 and compare with Table 1 and Appendix Table A1). A possible reason is that the economy slowed in 2009 after a period of rapid growth.
[21] In 2006, the mean utilization rate of the treatment group was 14.6%. The estimated effect of 0.049 is 34% of this.
[22] This is due to 2006 being an outlier year for the inpatient utilization rate in the comparison group. In 2005 and 2007 the rate is 7.6% but it dips to 5.7% in 2006. It is the recovery from this drop that is generating the negative point estimate despite a rise in the rate of inpatient utilization in the treatment group in the post-reform period (Table 2).

line. This is done to reduce the risk of selection bias due to healthcare utilization and related expenditure raising total household consumption above the poverty line (see footnote 12). Arguably, this leads to a discrepancy between our sample and the target population. And it may still result in selection on the outcome if medical expenses crowd-out expenditures on other goods. The results in column (3) of Table 3 reveal that, apart from the effect on utilization of diagnostic tests, the estimates are highly robust to sticking strictly to the official measure of poverty.

The results given in column (4) demonstrate that the estimates are generally robust to including the 880 poorest districts in which the anti-poverty strategy was implemented simultaneously to the extension of basic health insurance to poor adults. The estimated effect on diagnostic testing again falls to approximately half of the main estimate but in this case it remains marginally significant. The sensitivity of the estimated impact on this outcome is likely to due to the fact that it has the lowest prevalence – only 2.5% of the treatment group had a diagnostic test prior to the reform.

The estimates are generally robust not only to the selection of the sample but also to the estimator. The least squares estimate of the effect on ambulatory care given in the first column of Table 5 is closer to the simple DID estimate than it is to the logit estimate but it remains strongly significant. The estimated effect on medication is even more robust to estimation by least squares. The point estimate of the effect on diagnostic tests falls somewhat using least squares and, as might be expected, the standard error increases such that significance is lost.

In columns (2)-(4) of Table 5, the comparisons and with pre-reform treatment group observations are made observationally equivalent to the treated observations by applying inverse probability weights that are based on estimated propensity scores of being treated.[23] This is a flexible way of making the common trends identification assumption conditional on covariates that may be associated with the dynamics of the outcome (Abadie 2005). It is particularly advantageous in this application given the marked differences in covariates between the treatment and comparison groups. In addition to application of weights in columns (2)-(4), we also condition on covariates by including them in least squares regressions. This gives a doubly robust estimator that is consistent (given conditional independence) if either the propensity score model or the regression model, but not necessarily both, is correctly specified (Robins and Rotnitzky 1995). Besides

[23] The propensity score of being treated is estimated by three separate probit models that discriminate between treated observations and, in turn, comparison observations pre-reform (*k=1*), comparisons post-reform (*k=2*) and treatment group observations pre-reform (*k=3*). Let the estimated propensity score for an individual in group *k* be $\Phi(\mathbf{X}_{it}\hat{\gamma}_k)$. Then, that observation is given a weight equal to $\Phi(\mathbf{X}_{it}\hat{\gamma}_k) \big/ 1-\Phi(\mathbf{X}_{it}\hat{\gamma}_k)$. Greater weight is therefore given to comparison observations and pre-reform treatment group observations that are more similar to the post-reform treatment group observations with respect to covariates. The treated observations are given a weight of 1.

correcting for selection on observables in this unrestrictive way, taking the difference-in-differences still allows for differences in time invariant unobservables across the treatment and comparison groups. Bootstrap standard errors are used.

Table 5: Least squares estimated of effects on healthcare utilization

No weights	Inverse probability weights		
	No trimming	Trim controls with weight > 1% sum of weights	Trim observations with propensity score > 0.9
(1)	(2)	(3)	(4)
Ambulatory consultation			
0.046***	0.060**	0.071**	0.068***
(0.014)	(0.029)	(0.028)	(0.019)
Medicines			
0.037**	0.028	0.046	0.036*
(0.014)	(0.036)	(0.033)	(0.021)
Diagnostic test or examinations			
0.006	0.006	0.008	0.014
(0.007)	(0.010)	(0.010)	(0.012)
Visit to dentist or ophthalmologist			
0.001	0.050	0.009	-0.010
(0.009)	(0.044)	(0.015)	(0.013)
Hospitalized or surgery			
-0.007	0.046	0.004	-0.003
(0.007)	(0.046)	(0.020)	(0.011)
Sample size *53,713*	*53,713*	*53,712*	*38,246*

Notes: All models include the covariates listed in Table 1 and Appendix Table A1, plus 12 age-sex group indicators, 25 region indicators, eight housing density indicators, a post-treatment period indicator and month effects. Robust standard errors corrected for within household clustering reported in parenthesis in column (1) (Rogers 1994) and bootstrapped standard errors (1,000 repetitions) in columns (2)-(4). ***, **,* indicate statistical significance at 1, 5 and 10%.

In general, the estimates are reasonably robust to the application of inverse probability weights. In column (2), all observations are used. The effect on ambulatory care increases in magnitude relative to the unweighted OLS estimate and differs little from the logit estimate. The estimated effect on medication declines in magnitude and loses significance. Using all observations potentially makes the estimates highly sensitive to comparison observations that are given very large weights. To avoid this, column (3) trims the sample by dropping any observation (there is only one) with a weight that is more than one percent of the sum of all weights (Huber, Lechner, and Wunsch 2013). The estimated effect both on ambulatory care and on medication increases in magnitude, and the effect on the latter moves closer to significance (p-value=0.163). Extrapolation to treated observations that are off the support of the distribution of propensity scores for the comparisons may also be a concern when conditioning on covariates. Restriction to the common support comes at the cost of no longer estimating the average effect across all in the target treatment group. Nonetheless, in column (4) we drop all observations with an estimated propensity score greater than 0.9 (Crump et al. 2009). Despite the fact that this results in the loss of around 29% of the

sample, the estimated effect on ambulatory care differs little from that obtained with the one heavily weighted comparison observation dropped. The estimated effect on medication use is the same as the unweighted least squares estimate and similar to the logit estimate, and it is significant at the 10% level.

Overall, estimates obtained from various samples and estimators all point to the extension of coverage having raised utilization of ambulatory care and medication in the target group of poor adults, but having had no impact on their access to dental/ophthalmic and inpatient care, which the programme covered only to a limited degree. There is some evidence that utilization of diagnostic tests increased, but the estimated effect is less robust, probably because prevalence is very low.

4.5 Heterogeneous effects

The analysis conducted in the previous sub-section revealed some sensitivity of the estimates to narrowing the window of analysis around the reform. A possible reason is that the full effect of the coverage extension was not reached in the first year after implementation because beneficiaries were not yet aware of their newly acquired entitlement. We assess this hypothesis by using two years of data either side of the reform but allowing the effect to differ in the first and second post-reform years. Specifically, we extend (1) by interacting the treatment group indicator with a year 2008 dummy and a year 2009 dummy. Average partial effects of each interaction calculated over the treatment group sample observed in each of these years are given in the first two columns of Table 6. For all types of healthcare, the point estimate of the effect is larger in 2009 than it is in 2008, but none of the differences is statistically significant. The coverage extension is estimated to have raised the probability of consultation by 6.9 percentage points in 2009 and by 5.2 points in 2008. Since the nonlinearity of the logit model can produce differential effects, we also test for heterogeneity using the linear probability model. The results given in Appendix Table A5 are generally consistent with the logit estimates with respect to differences in the point estimates by year but again no difference is significant. There is some indication of a delay in the reform taking full effect but the evidence is not strong.

Within the adult population, healthcare needs rise strongly with age. If the poor were previously forgoing needed healthcare, then the extension of coverage should have a larger impact in the older population, at least in absolute terms. We test for this by splitting the sample into younger (20-49 years) and older (50+) individuals and estimating a set of models separately for each. The resulting average partial effects are presented in the third and fourth columns of Table 6. All point estimates are larger for the older sample. The differences are significant for ambulatory

consultations, diagnostic testing and dental/ophthalmic care. It is estimated that while the reform raised the probability of an ambulatory consultation in the older population by 8.8 percentage points, the impact on the younger adult population was 4.2 points. This difference does not simply reflect the elderly's higher rate of utilization. Relative to the baseline mean, the effect on older individuals is still approximately twice that on younger adults. The estimated effect on the probability of receiving medication is four-fifths larger for the older group – a 5.2 compared with a 2.9 percentage point increase. There is a significant impact on the likelihood of getting a diagnostic test only for older individuals. This effect is large: the utilization rate is estimated to almost double as a result of the reform. Even for dental or ophthalmic treatment there is a significant effect for the older group that is 39% of the baseline utilization rate. For inpatient treatment, the point estimate is positive only for the older sample but even then it is not remotely significant. Overall, the results suggest that the reform had a substantially larger impact on the healthcare utilization of older compared with younger poor adults. Using a linear probability model to estimate the effects does not change this conclusion (see Appendix Table A5).

If the removal of user fees raised utilization by making previously unaffordable care accessible, then the effect should be greater among the poorest. To test this, we split the sample by household per capita consumption below and above the median, and estimate effects for each group (Table 6). Bearing in mind that the sample consists of poor adults only, this discriminates between the poorest and the not-quite-so-poor. The estimated effect on ambulatory care is greater among the poorest (10% level of significance). The point estimate of the effect on medication is also larger for the poorest, but this difference is not significant. Moreover, there is a significant positive effect on the use of dental and ophthalmic care only for the poorest. In contrast, the estimated effect on diagnostic testing is larger and significant only for the less poor. Although this difference across the groups is not significant, it is consistent with diagnostic tests, which may not be fully covered, remaining unaffordable for the poorest. But for a general loss of precision, the ordering of effects by poverty severity remains intact using ordinary least squares (Appendix Table A5).

Given poverty is considerably more severe in rural areas where accessibility remained limited even after the affordability constraint on utilization was relaxed, any greater impact of the programme on those previously least able to afford healthcare could be partially or fully offset by a smaller effect in the rural localities where these individuals reside disproportionately.[24] In the estimation sample, 77% of urban dwellers but only 30% of rural dwellers lived in a district with a public hospital.[25] Only 3% of those in urban locations did not have a health centre within their district,

[24] In 2007, the poverty rate was 27% in urban areas and 66% in rural areas.
[25] Figures obtained by linking ENAHO with Peru's National Registry of Municipalities (RENAMU)

compared with 18% of those in rural areas. Granting entitlement to free healthcare may have little impact if medical facilities and/or manpower are insufficient to respond to the increased demand. The final two columns of Table 6 provide estimates for urban and rural dwellers separately.[26] The point estimates for ambulatory care, medication and diagnostic tests are all larger for the urban sample but none of the differences is significant. The absence of a clear discrepancy could be due to the offsetting effects of differential accessibility to facilities and severity of poverty.[27] In any case, there is no strong evidence the extension of coverage had no impact on utilization of healthcare in rural areas with a lower density of health facilities and manpower.

Table 6: Heterogeneous effects on healthcare utilization

	Post-reform year		Age		Non-health consumption		Area of residence	
	2008	2009	<50	≥50	≥ median	< median	Urban	Rural
Ambulatory consultation								
	0.052***	0.069***	0.042***	0.088***	0.055***	0.082***	0.064***	0.056***
	(0.012)	(0.012)	(0.013)	(0.015)	(0.012)	(0.015)	(0.012)	(0.017)
Baseline	0.145		0.139	0.161	0.159	0.133	0.154	0.134
p-value	0.318		0.038		0.078		0.911	
Medicines								
	0.034**	0.042**	0.029*	0.052**	0.030*	0.048*	0.039**	0.030
	(0.017)	(0.017)	(0.017)	(0.022)	(0.017)	(0.025)	(0.017)	(0.024)
Baseline	0.332		0.314	0.378	0.366	0.301	0.367	0.290
p-value	0.678		0.431		0.499		0.811	
Diagnostic test or examination								
	0.007*	0.009**	0.001	0.022***	0.010**	0.005	0.012***	0.004
	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)	(0.006)	(0.004)	(0.005)
Baseline	0.025		0.023	0.029	0.031	0.019	0.032	0.017
p-value	0.710		0.014		0.839		0.584	
Visit to dentist or ophthalmologist								
	-0.001	0.004	-0.011	0.018***	-0.003	0.014*	-0.001	0.008
	(0.007)	(0.007)	(0.008)	(0.006)	(0.008)	(0.008)	(0.008)	(0.009)
Baseline	0.063		0.069	0.046	0.078	0.049	0.077	0.046
p-value	0.557		0.003		0.113		0.437	
Hospitalized or surgery								
	-0.003	0.002	-0.002	0.004	-0.003	-0.001	-0.000	-0.003
	(0.006)	(0.006)	(0.006)	(0.008)	(0.006)	(0.010)	(0.006)	(0.009)
Baseline	0.036		0.035	0.038	0.040	0.032	0.042	0.029
p-value	0.526		0.564		0.908		0.797	
N	*53,713*		*37,592*	*16,121*	*26,858*	*26,855*	*29,529*	*24,184*

Notes: DID estimates from logit models. Year-specific average partial effects obtained from a model that allows the coefficient on the *Treat* × *Post* indicator to vary across the two years. Average partial effects for remaining sub-groups obtained from split-sample estimates. Treatment and comparison groups defined as in notes to Table 1. Covariates as in notes to Table 3. 'Baseline' is the mean outcome in the sub-treatment group in the pre-reform period. p-value is for Wald test of null of equal effects across years and sub-populations. Robust standard errors adjusted for clustering at the household level in parentheses (Korn and Graubard 2011). ***, ** and * indicates significance at 1%, 5% and 10% respectively.

5. Impact on healthcare expenditure

The stated objective of the 2007 reform was to increase utilization of healthcare by the poor and so reduce inequality in access. Pursuing this through targeted removal of user fees at public facilities promised a second potential benefit – reduced exposure to medical expenditure risk. In

[26] To ensure a sufficient number of comparison observations in the subsample of rural residents, we deviate from the official definition of a rural area by setting the threshold at less than 4000 homes, rather than less than 400 homes.

[27] Given the limited size of our comparison group, particularly in rural areas, it is not possible to estimate effects across groups discriminated by location and severity of poverty simultaneously.

this section, we examine whether the fee waiver succeeded in reducing OOP payments for healthcare incurred by the poor.

5.1 Outcomes and simple difference-in-differences

Out-of-pocket spending on each of 15 categories of healthcare and medicines is recorded for each person in the household.[28] Since poor children and pregnant women were covered prior to the reform and are dropped from the estimation sample, we exclude spending on child and maternity care. We also exclude expenditure on items that are not part of the SIS programme benefit package.[29] We aggregate spending over the remaining categories, after pro rata scaling to adjust for different recall periods, to get total OOP spending per person for a four week period.

Table 7 shows mean spending in the pre- and post-reform periods for both the treatment and comparison groups. Prior to the reform, the sample mean of spending in the treatment group was slightly greater than that of the comparison group but the difference is not significant (top panel). This may seem surprising given that only the comparison group had insurance in this period. One must keep in mind that the substantially lower utilization of the treatment group (Table 2) offsets the higher price it faced. In the bottom panel, we restrict attention to individuals who used formal healthcare and/or received medication. In this sample of patients, mean spending prior to the reform was significantly higher in the treatment group than in the comparison group (p-value=0.000). Those who were using healthcare without insurance were paying more out-of-pocket than those who had insurance.

In the full sample, mean spending increased for both groups between the pre- and post-reform periods. The change is smaller in the treatment group that acquired insurance cover, which results in a negative simple DID estimate that is not significant. Again, one must factor in the greater increase in utilization of the treatment group (at least of ambulatory care and medication). Restricting attention to those who used healthcare, the point DID estimate is again negative and is substantially larger in magnitude than that obtained from the full sample but it is still not remotely significant. This simple analysis produces no evidence to suggest that the removal of user fees for poor adults reduced their OOP expenditures.[30]

[28] We include all OOP spending on treatment for someone within the household irrespective of whether the payment was made by the household itself or by another on its behalf.

[29] Specifically, spectacles, contraceptives and orthopaedic and other devices. Spending on vaccinations is also excluded since these are provided by vertical programmes.

[30] To gauge the plausibility of the identification assumption on which the simple DID estimates rest, as for the utilization outcomes, we test whether the trends in the pre-treatment period were equal between the treatment and comparison groups. The results in Appendix Table A6 (columns indicating no control for covariates) show that the null of equal trends is never rejected.

Table 7: Out-of-pocket expenditure means (Soles Nuevos)

	Pre-reform (2005/6)	Post-reform (2008/9)	Difference (post-pre)	Difference-in-differences	*Sample size (by group)*
Full sample					
Treatment	10.618	12.772	2.154***	-0.750	*47,284*
	(0.270)	(0.393)	(0.476)	(1.622)	
Comparison	9.963	12.867	2.905*		*5,528*
	(0.679)	(1.395)	(1.551)		
Sample size (by period)	*30,866*	*21,946*			*52,812*
If use formal healthcare or medicines					
Treatment	27.708	28.761	1.053	-3.479	*19,320*
	(0.671)	(0.854)	(1.086)	(3.031)	
Comparison	19.201	23.733	4.532		*2,920*
	(1.269)	(2.531)	(2.831)		
Sample size (by period)	*12,277*	*9,963*			*22,240*

Notes: Amounts are per person for a four-week period. The sample, as well as the treatment and comparison groups, defined as in notes to Table 1. OOP expenditure amounts in 2006 Soles Nuevos (S/.), with 1S./≈0.3$US. 'Difference-in-differences' is the post-pre reform difference in means of the treated minus that of the comparisons. Robust standard errors adjusted for clustering within households in parentheses (Rogers 1994). ***, **,* indicate significantly different from zero at 1%, 5% and 10%.

5.2 Conditional difference-in-differences

We condition on observables using a Generalised Linear Model (GLM) that specifies mean medical expenditure (m_{it}) as an exponential function of covariates,

$$E\left[m_{it} \left| Treat_i, Year_t, \mathbf{X}_{it} \right. \right] = \exp\left(\theta Treat_i + \tau_t Year_t + \delta Treat_i \times Post_t + \mathbf{X_{it}}\boldsymbol{\beta} \right) \tag{3}$$

where we use the same notation as in (1) but obviously the parameters are different. We estimate the conditional means by Gamma Pseudo Maximum Likelihood (GPML) (Gourieroux, Monfort, and Trognon 1984; Manning and Mullahy 2001; Silva and Tenreyro 2006). This estimator has been shown to perform well even when there is a large proportion of observations with zero values (Silva and Tenreyro 2011). In our sample, 63% of individuals report no OOP spending.

Under the assumption that the proportionate change in spending would have been the same in the treatment and comparison groups if there had been no reform, the interaction term in (3) is attributable entirely to the policy and the intention-to-treat effect is (Lechner 2011):

$$ITT^m = \frac{1}{N_{TP}} \sum_i 1\left(Treat_i \times Post_t = 1\right) \exp\left(\hat{\theta} + \hat{\tau}_t Year_t + \mathbf{X_{it}}\hat{\boldsymbol{\beta}} \right) \left[\exp\left(\hat{\delta} \right) - 1 \right] \tag{4}$$

The term in square brackets is the effect relative to counterfactual expenditure of the treated in the absence of the reform.

The top panel of Table 8 gives estimates obtained from the full sample. The point estimates in column (1) are negative, but not remotely significant. If we select the sample using the official measure of poverty, then the estimated negative impact of the fee-waiver on the level of OOP

spending is borderline significant (p-value=0.108) and the relative effect indicates a reduction of 18%. If the sample is extended to include the poorest districts in which the anti-poverty strategy was implemented simultaneous to the extension of health insurance to poor adults, then the absolute effect increases in magnitude and reaches significance at the 10% level, while the relative reduction increases to 20%. But we are apprehensive of attributing this to the coverage extension alone.

Table 8: Estimated effects on out-of-pocket expenditure

	Main estimate	Robustness to using		
		only 2006 (pre) & 2008 (post) data	official measure to select poor	all districts, incl. 880 poorest
	(1)	(2)	(3)	(4)
Full sample				
ITT	-1.249	0.117	-1.562	-2.587*
	(1.511)	(1.841)	(0.971)	(1.354)
Relative effect	-0.090	0.009	-0.175*	-0.203**
	(0.099)	(0.142)	(0.090)	(0.086)
Sample size	*52,812*	*26,405*	*45,501*	*87,679*
If use formal healthcare or medicines				
ITT	-5.101	-0.362	-5.068**	-8.836***
	(3.339)	(3.871)	(2.222)	(3.320)
Relative effect	-0.154*	-0.012	-0.224***	-0.268***
	(0.087)	(0.130)	(0.077)	(0.074)
Sample size	*22,240*	*11,357*	*17,759*	*34,003*

Notes: ITT are GPML estimates obtained from equation (4). They are effects on individual spending over four week period (Soles Nuevos). Relative effect is $\exp(\hat{\delta})-1$ from equation (4). The sample, as well as the treatment and comparison groups, defined as in notes to Table 1. Covariates as in notes to Table 3. Bottom panel restricts sample to individuals who have used formal healthcare or medicines. Robust standard errors adjusted for clustering within households in parentheses (Korn and Graubard 2011). ***, **,* indicate significantly different from zero at 1%, 5% and 10%.

In the bottom panel, we restrict the sample to individuals who made some use of formal healthcare or medicines. The main estimate of the ITT increases in magnitude and moves closer to significance (p-value=0.133), while the estimate of the relative effect indicates a significant 15% reduction in the amount patients spend out-of-pocket. If we use the official measure of poverty to select the sample, or include the poorest districts, then both the absolute and relative effects are significant.

Using the baseline selection of the sample, least squares estimates of the impact on the OOP amount are broadly consistent in magnitude and (non) significance with the GPML estimates given in column (1) of Table 8 (see Appendix Table A7, top panel).[31] Overall, while the estimates in

[31] Correcting for selection bias when restricting the sample to users of healthcare has little impact on the least squares estimates (see Appendix Table A7). Using least squares, the effect is identified under the assumption of parallel trends in the level of expenditure, as opposed to equi-proportionate changes in expenditure. The latter is again the identification condition if we model the log of expenditures by least squares. We do this for positive OOP payments in the bottom panel of Table A7. This also indicates no significant impact.

columns (3) and (4) of Table 8 hint at the possibility that extending cover to poor adults succeeded in reducing the payments they incurred when using healthcare, there is no robust evidence of an impact on mean OOP expenditure.

5.3 Effects across the distribution healthcare expenditure

Lack of strong evidence of an impact on the conditional mean does not rule out effects on other parameters of the distribution. Given the limited benefit package offered and the ceiling on coverage, the healthcare for the poor programme might be expected to have had little impact on extremely high expenses that exert a strong influence on the mean of the typically skewed distribution of medical expenditure. The removal of user fees for basic services at public facilities could, however, reduce more modest levels of spending.

We estimate censored quantile regressions (CQR) that take account of the large fraction of observations with no OOP payments (Chernozhukov and Hong 2002).[32] The conditional quantile function is specified as,

$$Q_{m|.}(q) = \max\left(\theta_q Treat_i + \tau_{tq} Year_t + \delta_q Treat_i \times Post_t + \mathbf{X_{it}}\boldsymbol{\beta}_q, 0\right). \quad (5)$$

The difference-in-differences between the treatment and comparison groups in the (conditional) quantile *q* of the uncensored distribution is δ_q. If without the extension of coverage, this quantile of the medical spending of treated individuals would have changed as did quantile *q* of the spending of comparison households, then δ_q corresponds to the effect of the reform on that (conditional) quantile (Athey and Imbens 2006). We estimate the average *corner adjusted* effect by multiplying the CQR estimate of δ_q by the proportion of households with positive predicted expenditure at the conditional quantile (Limwattananon et al 2015; Kowalski 2016). We estimate the effect relative to counterfactual spending at the quantile by the median of $\frac{\hat{\delta}_q}{\hat{\theta}_q + \hat{\tau}_{tq} Year_t + \mathbf{X_{it}}\hat{\boldsymbol{\beta}}_\mathbf{q}}$ over treated observations with a positive predicted quantile. We report 90% confidence intervals constructed from a bootstrap.[33]

The top panel of Figure 3 shows the estimates of effects from the 50th to the 90th quantiles in steps of 5 that are obtained from the full sample.[34] The absolute and relative effects are presented in the

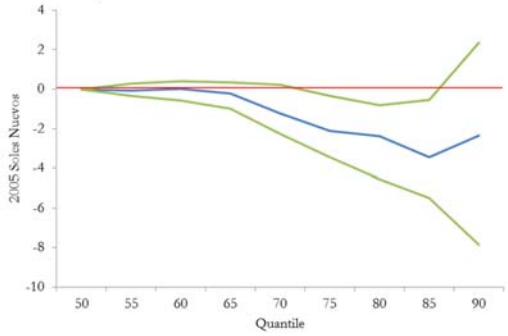
[32] Computation is done using an adaptation to the Stata® ado cqiv (Chernozhukov et al. 2012). We are grateful to the programmers of the ado for making it available.
[33] We draw 499 bootstrap samples with replacement.
[34] Because of the large proportion of zeros in the sample, most conditional quantiles and estimated effects below the 50th quantile are zero. Above the 90th quantile, the confidence interval is very wide. See Appendix Table A8 for the estimates.

left-hand and right-hand columns respectively. The point estimate of the absolute effect is very close to zero between the 50th and 65th quantile and is not significant below the 75th quantile. There is a significant reduction in OOP expenditure that increases in magnitude from the 75th to the 85th quantiles. The increasing magnitude is simply an artefact of moving up the distribution. The relative effect declines from a 26% reduction at the 75th quantile to a 17% decrease at the 85th quantile.

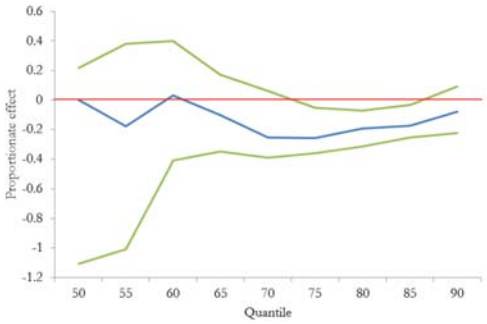
Figure 3: Effects on conditional quantiles of out-of-pocket expenditure

Full sample

a) Absolute effect

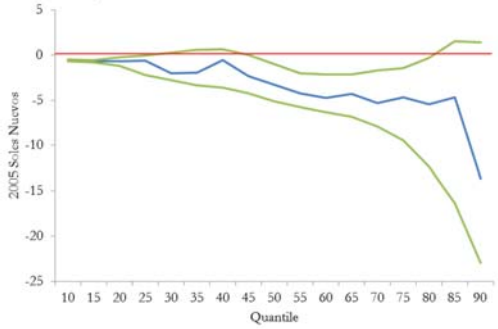


b) Relative effect

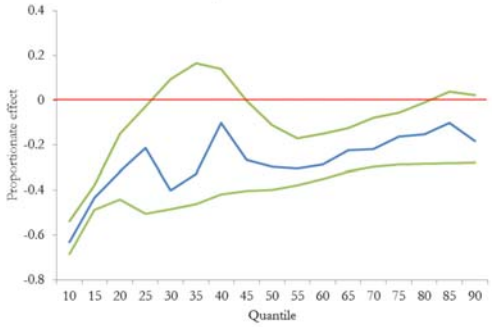


If use formal healthcare or medicines

a) Absolute effect



b) Relative effect



Notes: Effects estimated from censored (at zero) quantile regressions in blue. Bootstrapped 90% confidence intervals in green.

In the bottom panel we show the effects in the sample of individuals who make use of formal healthcare or medicines. In this sample there are far fewer with zero OOP payments and we show the effects from the 10th to the 90th quantile. Spending is reduced significantly at the bottom of the distribution. In relative terms, the reform is estimated to have cut OOP expenditure at the 10th conditional quantile by 63%. The effect decreases in relative terms but remains significant until the 25th quantile, at which point it is 21%. The effect is imprecisely estimated from the 25th to the 45th quantiles. Thereafter, it is significant and decreasing in relative terms until the 80th quantile, where it is 15%.

These results suggest that the programme may well have reduced OOP expenses. This is not apparent in mean spending of the full sample because the positive impact on utilization of services that are not fully covered offsets the reduction in price for a given quantity, and also because the effects differ across the distribution. Consistent with the narrow benefit package offered, the impact is strongest in relative terms on lower levels of spending of healthcare users.

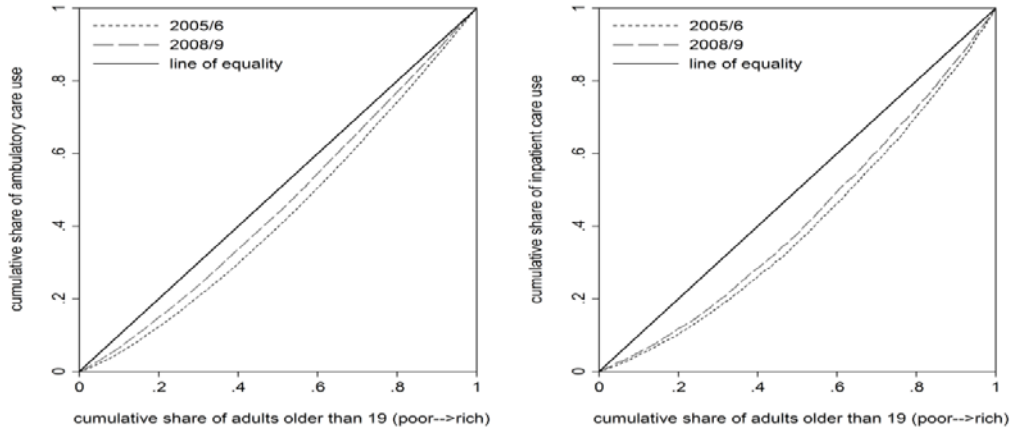
Conclusion

Health systems of low- and middle-income countries that give priority to providing the poor with basic healthcare aim to correct the most blatant inequities. Achieving this is not a foregone conclusion. There are plenty examples of subsidies that are inefficiently targeted on the poor, fee waivers that are ineffectively implemented and public health facilities that are so lacking in manpower and supplies that even the poor are discouraged from using them. Before *progressive universalism* (Gwatkin and Ergo, 2011; Jamison et al, 2013) becomes a shallow slogan that is hijacked in defence of underperforming public health systems, it is important to establish that targeting can be effective in improving the poor's access to healthcare and their protection from medical expenditure risks.

This paper demonstrates that granting poor Peruvians entitlement to free basic healthcare has been partially successful in realizing these goals. We estimate that it raised the one-month utilization rate of ambulatory care by six percentage points against a baseline of 14.5. This very substantial increase in access to first line medical treatment was accompanied by a one third increase in the rate of diagnostic testing, while the medication rate was raised more modestly by 11%. These gains appear to have contributed to a reduction in inequality in utilization of ambulatory care that is evident in the significant inward shift of the concentration curve on the left-hand panel of Figures 4a.[35] The figure on the right reveals that the substantially greater inequality in access to inpatient treatment also decreased between the pre- and post-reform periods. But consistent with our finding of no significant impact of the coverage extension on inpatient care, the shift in this distribution is appreciably more modest. This reflects the design of the policy, which offered a narrow benefit package and imposed ceilings on coverage. It is also likely due to the fact that reimbursement only for medication and consumables but not salary costs gave facilities very little incentive to deliver inpatient treatment even for conditions that were in the benefit package. And it likely reflects lack of physical access of the poor rural population to distant hospitals providing inpatient treatment.

[35] In both figures, the 2005-06 curve dominates (lies significantly below) the 2008-09 according to the test of Bishop, Chow, and Formby (1994).

Figure 4: Concentration curves for ambulatory care (left) and inpatient care (right) in pre-reform (2005/6) and post-reform (2008/9) periods.



Notes: Sample as described in notes to Table 1. Individuals are ranked by per capita household consumption net of medical expenditure. Both healthcare indicators are binary.

The lack of any significant impact on mean out-of-pocket spending on healthcare may be attributable to both the lack of effective coverage of expensive hospital care and the offsetting effect of increased utilization of only partially covered treatments. Restricting attention to users of healthcare, we estimate a 15% drop in mean payments and the proportionate reduction in expenses incurred by patients is even larger below the bottom quartile of the distribution. The smaller relative impact further up the distribution again reflects the limited benefit package. Basic insurance does not protect against major medical expenditure risks. With limited public funds, reducing exposure to such risks would require cutting back on first dollar coverage of basic healthcare. The goal of the reform in Peru was to improve the poor's access to healthcare. Against this objective, increased utilization of ambulatory care and medicines along with a more modest impact on out-of-pocket spending may still be judged a success.

In any case, when insurance raises utilization, its impact on financial risks associated with illness cannot be inferred only from the effect on OOP expenditure. Spending would have fallen by more if utilization had remained constant. If utilization of treatments that are only partially covered rises sufficiently, then the impact on spending unconditional on use could even be positive. This is the explanation offered by Bernal et al (2015) for their positive estimated effect on OOP expenditure of a later version of the health insurance for the poor programme in Lima. This is a more credible scenario in the capital city, where there is an abundance of private providers ready to meet inflated demand, than it is in the periphery. The reason we do not find this paradoxical result may be because we estimate the effect of a substantially more limited programme in the whole country. It may also be because we estimate the impact across the target poor population, as opposed to the effect on those at the margin of poverty. The latter might be able to afford partially, or even wholly

uncovered, treatments as a result of increased awareness of their medical care needs, while those in deeper poverty may have to settle for what they get for free.

References

Abadie, Alberto. 2005. "Semiparametric Difference-in-Differences Estimators." *The Review of Economic Studies* 72 (1): 1-19.

Arróspide, M., K. Rozas, and J. Valderrama. 2009. *Presupuesto Evaluado: Seguro Integral De Salud*. Lima, Peru: Ministerio de Economía y Finanzas.

Athey, Susan and Guido W. Imbens. 2006. "Identification and Inference in Nonlinear Difference□ in□ differences Models." *Econometrica* 74 (2): 431-497.

Bauhoff, Sebastian, David R. Hotchkiss, and Owen Smith. 2011. "The Impact of Medical Insurance for the Poor in Georgia: A Regression Discontinuity Approach." *Health Economics* 20 (11): 1362-1378.

Bernal, Noelia, Miguel A. Carpio, and Tobias J. Klein. 2015. "The Effects of Access to Health Insurance for Informally Employed Individuals in Peru." *Mimeo University of Piura, Lima*.

Bishop, John A., K. Victor Chow, and John P. Formby. 1994. "Testing for Marginal Changes in Income Distributions with Lorenz and Concentration Curves." *International Economic Review*: 479-488.

Borra, Virginia. 2010. *Estrategia Nacional CRECER - Política De Gestión De La Secretaría Técnica De La Comisión Interministerial De Asuntos Sociales*. Lima: Comisión interministerial de asuntos sociales.

Chernozhukov, Victor, Ivan Fernandez-Val, Sukjin Han, and Amanda Kowalski. 2012. "CQIV: Stata Module to Perform Censored Quantile Instrumental Variables Regression." *Statistical Software Components*.

Chernozhukov, Victor and Han Hong. 2002. "Three-Step Censored Quantile Regression and Extramarital Affairs." *Journal of the American Statistical Association* 97 (459).

Cotlear, Daniel, Somil Nagpal, Owen Smith, Ajay Tandon, and Rafael Cortez. 2015. *Going Universal: How 24 Developing Countries are Implementing Universal Health Coverage from the Bottom Up*. Washington, D.C.: World Bank Publications.

Crump, Richard K., V. Joseph Hotz, Guido W. Imbens, and Oscar A. Mitnik. 2009. "Dealing with Limited Overlap in Estimation of Average Treatment Effects." *Biometrika*: asn055.

Diaz, Juan José and Martin Valdivia. 2011. "The Vulnerability of the Uninsured to Health Shocks in Peru." In *Household Spending and Impoverishment. Volume 1 of Financing Health in Latin America Series.*, edited by Felicia Marie Knaul, Rebeca Wong and Héctor Arreola-Ornelas. Vol. 53, 249-285. Cambridge, MA: Cambridge, MA: Harvard Global Equity Initiative, in collaboration with Mexican Health Foundation and International Development Research Centre, distributed by Harvard University Press.

EsSalud. 2010. *Memoria Anual 2009*. Lima, Peru: EsSalud.

———. . 2006. *Memoria Institucional 2005*. Lima, Peru: EsSalud.

———. . 2007. *Memoria Institucional 2006*. Lima, Peru: EsSalud.

———. . 2008. *Memoria Institucional 2007*. Lima, Peru: EsSalud.

Flores, Gabriela, Por Ir, Chean R. Men, Owen O'Donnell, and Eddy Van Doorslaer. 2013. "Financial Protection of Patients through Compensation of Providers: The Impact of Health Equity Funds in Cambodia." *Journal of Health Economics* 32 (6): 1180-1193.

Francke, Pedro. 2013. *Peru's Comprehensive Health Insurance and New Challenges for Universal Coverage. Universal Health Coverage (UNICO) Studies Series; no. 11*. Washington, DC: World Bank Group.

Gertler, Paul, Luis Locay, and Warren Sanderson. 1987. "Are User Fees Regressive?: The Welfare Implications of Health Care Financing Proposals in Peru." *Journal of Econometrics* 36 (1): 67-88.

Gourieroux, Christian, Alain Monfort, and Alain Trognon. 1984. "Pseudo Maximum Likelihood Methods: Applications to Poisson Models." *Econometrica: Journal of the Econometric Society*: 701-720.

Grogger, J., T. Arnold, A. S. Leon, and A. Ome. 2015. "Heterogeneity in the Effect of Public Health Insurance on Catastrophic Out-of-Pocket Health Expenditures: The Case of Mexico." *Health Policy and Planning* 30 (5): 593-599.

Gwatkin, Davidson R. and Alex Ergo. 2011. "Universal Health Coverage: Friend Or Foe of Health Equity?" *The Lancet* 377 (9784): 2160-2161.

Huber, Martin, Michael Lechner, and Conny Wunsch. 2013. "The Performance of Estimators Based on the Propensity Score." *Journal of Econometrics* 175 (1): 1-21.

INEI. 2006. *Encuesta Nacional De Hogares - ENAHO 2006*. Lima, Perú: Instituto Nacional de Estadística e Informática.

Jamison, Dean T., Lawrence H. Summers, George Alleyne, Kenneth J. Arrow, Seth Berkley, Agnes Binagwaho, Flavia Bustreo, et al. 2013. "Global Health 2035: A World Converging within a Generation." *The Lancet* 382 (9908): 1898-1955.

King, Gary, Emmanuela Gakidou, Kosuke Imai, Jason Lakin, Ryan T. Moore, Clayton Nall, Nirmala Ravishankar, Manett Vargas, Martha María Téllez-Rojo, and Juan Eugenio Hernández Ávila. 2009. "Public Policy for the Poor? A Randomised Assessment of the Mexican Universal Health Insurance Programme." *The Lancet* 373 (9673): 1447-1454.

Korn, Edward L. and Barry I. Graubard. 2011. *Analysis of Health Surveys*. Vol. 323 John Wiley & Sons.

Kowalski, Amanda. 2016. "Censored Quantile Instrumental Variable Estimates of the Price Elasticity of Expenditure on Medical Care." *Journal of Business & Economic Statistics* 34 (1): 107-117.

Lagarde, Mylene and Natasha Palmer. 2008. "The Impact of User Fees on Health Service Utilization in Low-and Middle-Income Countries: How Strong is the Evidence?" *Bulletin of the World Health Organization* 86 (11): 839-848C.

Lechner, Michael. 2011. *The Estimation of Causal Effects by Difference-in-Difference Methods*. Foundations and Trends in Econometrics: Vol. 4: No. 3, pp.165-224 ed. Boston - Delft: Now publishers inc.

Lépine, A., M. Lagarde, and A. Le Nestour. 2015. "Free Primary Care in Zambia: An Impact Evaluation using a Pooled Synthetic Control Method." *Available at SSRN:Http://Ssrn.Com/abstract=2520345*.

Limwattananon, Supon, Sven Neelsen, Owen O'Donnell, Phusit Prakongsai, Viroj Tangcharoensathien, Eddy van Doorslaer, and Vuthiphan Vongmongkol. 2015. "Universal Coverage with Supply-Side Reform: The Impact on Medical Expenditure Risk and Utilization in Thailand." *Journal of Public Economics* 121: 79-94.

Manning, Willard G. and John Mullahy. 2001. "Estimating Log Models: To Transform Or Not to Transform?" *Journal of Health Economics* 20 (4): 461-494.

Miller, G., D. Pinto, and M. Vera-Hernández. 2013. "Risk Protection, Service use and Health Outcomes Under Colombia's Health Insurance Program for the Poor." *American Economic Journal: Applied Economics* 5 (4): 61-91.

Ministerio de Salud. 2015. *Personal Del Ministerio De Salud Y Gobiernos Regionales Por Grupos Ocupacionales*.

Nyman, John A. 1999. "The Value of Health Insurance: The Access Motive." *Journal of Health Economics* 18 (2): 141-152.

Powell-Jackson, Timothy, Kara Hanson, Christopher JM Whitty, and Evelyn K. Ansah. 2014. "Who Benefits from Free Healthcare? Evidence from a Randomized Experiment in Ghana." *Journal of Development Economics* 107: 305-319.

Puhani, Patrick A. 2012. "The Treatment Effect, the Cross Difference, and the Interaction Term in Nonlinear "difference-in-Differences" Models." *Economics Letters* 115 (1): 85-87.

Robins, James M. and Andrea Rotnitzky. 1995. "Semiparametric Efficiency in Multivariate Regression Models with Missing Data." *Journal of the American Statistical Association* 90 (429): 122-129.

Rogers, William. 1994. "Regression Standard Errors in Clustered Samples." *Stata Technical Bulletin* 3 (13).

Silva, JMC Santos and Silvana Tenreyro. 2011. "Further Simulation Evidence on the Performance of the Poisson Pseudo-Maximum Likelihood Estimator." *Economics Letters* 112 (2): 220-222.

———. 2006. "The Log of Gravity." *The Review of Economics and Statistics* 88 (4): 641-658.

SIS. 2007. *Memoria Anual Del Seguro Integral De Salud*. Lima, Perú: Seguro Integral de Salud.

———. . 2011. *Memoria Institutional Año 2010
*. Lima, Perú: Seguro Integral de Salud.

Sparrow, Robert, Asep Suryahadi, and Wenefrida Widyanti. 2013. "Social Health Insurance for the Poor: Targeting and Impact of Indonesia's Askeskin Programme." *Social Science & Medicine* 96: 264-271.

Vermeersch, Christel, Andre C. Medici, and Rory Narvaez. 2014. *Peru - Universal Health Coverage for Inclusive and Sustainable Development : Country Summary Report*. Washington, DC: World Bank Group.

Wagstaff, Adam. 2010. "Estimating Health Insurance Impacts Under Unobserved Heterogeneity: The Case of Vietnam's Health Care Fund for the Poor." *Health Economics* 19 (2): 189-208.

World Bank. 2011. *Peru Recurso Programmatic AAA - Phase IV Improving Health Outcomes by Strengthening Users' Entitlements and Reinforcing Public Sector Management*. Washington DC: World Bank Group.

World Health Organization. "Global Health Expenditure Database.", accessed 2/6, 2016, http://apps.who.int/nha/database.

———. 2013. "A Universal Truth: No Health without a Workforce." *2013-2-24].Http://Www.Who.Int/Workforcealliance/Knowledge/Resources/hrhreport2013/En/Index.Html*.

Appendix Tables

Table A1: Covariate means and changes in means by treatment status, additional covariates

	Pre-reform (2005/6)			Change pre-post reform (2005/6-2008/9)		
	Treatment (T)	Comparison (C)	T=C p-value	Treatment (△T)	Comparison (△C)	△T =△C p-value
Water source						
Tap in home	0.562	0.728	0.000	-0.058***	-0.020	0.024
Tap in building	0.038	0.039	0.391	0.034***	0.021***	0.112
Public fountain	0.034	0.033	0.821	0.004	-0.002	0.337
Water truck	0.034	0.028	0.074	-0.005*	0.011*	0.015
Well	0.079	0.053	0.000	-0.013***	-0.019**	0.458
River or canal	0.191	0.072	0.000	0.021***	-0.005	0.012
Other	0.062	0.046	0.000	0.017***	0.013*	0.666
Sewage type						
Public system in home	0.385	0.631	0.000	-0.042***	-0.020	0.221
Public system in building	0.034	0.040	0.037	0.018***	0.017**	0.818
Septic tank	0.209	0.136	0.000	-0.035***	-0.052***	0.138
None	0.372	0.192	0.000	0.058***	0.055***	0.856
Main cooking fuel						
Electricity or gas	0.303	0.595	0.000	0.014***	0.023	0.637
Kerosene	0.067	0.068	0.767	-0.046***	-0.044***	0.815
Charcoal	0.058	0.064	0.184	-0.001	-0.006	0.626
Wood	0.478	0.228	0.000	0.005	0.014	0.582
Other	0.075	0.028	0.000	0.033***	0.013*	0.006
Does not cook	0.018	0.017	0.471	-0.005***	0.001	0.195
Number of observations	*3,470*	*28,073*				

Notes: Sample of poor individuals aged 20+ who do not report use of maternity care in past 12 months and reside outside the 880 poorest districts. Treatment group (T) consists of those reporting cover through health insurance for the poor or no insurance. Comparison group (C) includes those reporting employment-based health insurance. ***, **, * indicate that the change over time is significant at 1%, 5% and 10% respectively. Means for other covariates included in the estimation models given in Table 1.

Table A2: Pre-reform trends in healthcare utilization

	Ambulatory consultation in last 4 weeks		Medicines in last 4 weeks		Diagnostic tests or examinations in last 4 weeks		Visit to dentist or ophthalmologist in last 3 months		Inpatient stay or surgery in last 12 months	
Treatment × 2005	-0.012	-0.129	-0.022	-0.088	0.011	0.061	0.027**	0.231**	-0.002	-0.175
	(0.016)	(0.080)	(0.018)	(0.073)	(0.009)	(0.139)	(0.011)	(0.112)	(0.009)	(0.132)
Treatment × 2006	0.011	-0.016	0.013	0.054	0.007	-0.098	0.003	0.074	0.018**	0.160
	(0.016)	(0.083)	(0.018)	(0.075)	(0.010)	(0.144)	(0.012)	(0.110)	(0.009)	(0.145)
Model	OLS	Logit	OLS	Logit	OLS	Logit	OLS	Logit	OLS	Logit
Covariates	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Sample size	*49,586*		*49,586*		*49,585*		*49,586*		*49,586*	

Notes: Treatment and comparison group defined as in notes to Table A1. Data from 2004, 2005 and 2006 ENAHO. Coefficients are OLS and logit estimates of treatment group × year interactions. All models include indicators for treatment group, year 2005 and year 2006. Logit estimates additionally control for covariates defined as in notes to Table 3 (except for cooking fuel, which is available only from July 2004). Robust standard errors robust to clustering at the household level in parentheses. ***, **,* indicate significantly different from zero at 1%, 5% and 10%.

Table A3: 'Effects' of placebo reform purportedly implemented in 2005

	Ambulatory consultation in last 4 weeks	Medicines in last 4 weeks	Diagnostic tests or examinations in last 4 weeks	Visit to dentist or ophthalmologist in last 3 months	Inpatient stay or surgery in last 12 months
'Effect'	-0.002	0.012	-0.002	0.005	0.005
	(0.010)	(0.016)	(0.004)	(0.006)	(0.004)
Baseline (treated)	0.166	0.348	0.030	0.064	0.048
Sample	*33,249*	*33,249*	*33,248*	*33,249*	*33,249*

Notes: Data from 2004 and 2006 ENAHO. 'Effect' is average partial effect of treatment group × year 2006 interaction derived from a logit model akin to equation (1). This is equivalent to a logit DID estimate of the placebo reform purportedly implemented in 2005. Treatment and comparison group defined as in notes to Table A1. Covariates as in notes to Table 3. 'Baseline (treated)' is the respective outcome variable's treatment group mean in 2004. Robust standard errors corrected for within household clustering in parentheses (Korn and Graubard 2011).

Table A4: Covariate means by treatment status, 2006-8 sample

	Pre-reform (2006)			Change pre-post reform (2006-08)		
	T	C	T=C p-value	△T	△C	△T =△C p-value
Demographics						
Age in years	41.261	46.630	0.000	1.276***	1.663**	0.632
Male	0.502	0.534	0.013	-0.005	0.011	0.263
Married or legal partnership	0.656	0.743	0.000	0.001	0.016	0.434
# of household members	5.611	5.989	0.000	-0.162***	-0.248*	0.516
Highest level of education						
None	0.105	0.058	0.000	0.016***	0.030***	0.232
Incomplete primary	0.228	0.163	0.000	0.012**	0.017	0.773
Complete primary	0.185	0.139	0.000	-0.011**	0.005	0.269
Incomplete secondary	0.162	0.110	0.000	-0.007	0.015	0.115
Complete secondary	0.220	0.276	0.000	-0.007	-0.013	0.744
Any tertiary	0.100	0.254	0.000	-0.003	-0.054***	0.004
Occupation						
Not working	0.218	0.306	0.000	-0.019***	-0.011	0.669
Agriculture/fishing	0.368	0.130	0.000	0.001	0.006	0.743
Mining/manufacturing/construction	0.105	0.151	0.000	0.017***	-0.005	0.154
Services	0.310	0.413	0.000	0.002	0.009	0.723
Job type						
Professional	0.020	0.117	0.000	-0.002	-0.021*	0.133
Clerical support, service, or sales worker	0.092	0.153	0.000	-0.008**	-0.004	0.764
Skilled agricultural/forestry/fishery worker	0.175	0.061	0.000	0.006	0.014	0.477
Craft and related trades worker	0.062	0.084	0.001	0.009**	0.008	0.918
Plant/machine operator, assembler	0.057	0.056	0.788	0.011***	-0.000	0.205
Elementary	0.375	0.224	0.000	0.003	0.014	0.521
Water source						
Tap in home	0.561	0.729	0.000	-0.059***	-0.024	0.128
Tap in building	0.038	0.044	0.224	0.040***	0.013	0.014
Public fountain	0.036	0.027	0.071	-0.000	0.006	0.498
Water truck	0.034	0.028	0.183	-0.001	0.012	0.156
Well	0.077	0.049	0.000	-0.012**	-0.014	0.860
River or canal	0.197	0.080	0.000	0.005	-0.007	0.397
Other	0.057	0.043	0.015	0.028***	0.014	0.211
Sewage type						
Public system in home	0.385	0.644	0.000	-0.042***	-0.043*	0.972
Public system in building	0.034	0.045	0.024	0.020***	0.010	0.346
Septic tank	0.231	0.152	0.000	-0.063***	-0.067***	0.829
None	0.351	0.160	0.000	0.085***	0.100***	0.500
Main cooking fuel						
Electricity or gas	0.319	0.628	0.000	-0.001	-0.028	0.272
Kerosene	0.051	0.052	0.868	-0.024***	-0.029***	0.599
Charcoal	0.056	0.062	0.363	-0.000	0.002	0.864
Wood	0.473	0.208	0.000	0.008	0.036*	0.206
Other	0.082	0.032	0.000	0.022***	0.014	0.475
Does not cook	0.019	0.017	0.553	-0.004*	0.006	0.125
Log monthly non-health consumption p.c.	4.898	5.169	0.000	0.050***	0.021	0.087
Living place with more than 400 homes	0.609	0.866	0.000	0.012	-0.007	0.332
N	*13,570*	*1,636*				

Notes: Pre-reform data from 2006, and post-reform data from 2008 ENAHO surveys. Sample selected and treatment (T) and comparison (C) groups defined as in Table A1. ***, ** and * indicate that the change over time is statistically significance at 1%, 5% and 10% respectively.

Table A5: Heterogeneous effects on healthcare utilization, least squares estimates

	Post-reform year		Age		Non-health consumption		Area of residence	
	2008	2009	<50	≥50	≥ median	< median	Urban	Rural
Ambulatory consultation								
	0.038**	0.055***	0.027	0.083***	0.038**	0.088***	0.053***	0.041
	(0.017)	(0.017)	(0.016)	(0.022)	(0.016)	(0.027)	(0.016)	(0.026)
Baseline	0.145		0.139	0.161	0.159	0.133	0.154	0.134
p-value	0.415		0.034		0.115		0.695	
Medicines								
	0.032*	0.042**	0.027	0.051**	0.029*	0.047*	0.039**	0.025
	(0.018)	(0.019)	(0.018)	(0.023)	(0.017)	(0.028)	(0.017)	(0.027)
Baseline	0.332		0.314	0.378	0.366	0.301	0.367	0.290
p-value	0.653		0.404		0.579		0.661	
Diagnostic test or examination								
	0.004	0.008	-0.003	0.025*	0.008	0.005	0.014	-0.014
	(0.009)	(0.010)	(0.008)	(0.014)	(0.009)	(0.015)	(0.009)	(0.014)
Baseline	0.025		0.023	0.029	0.031	0.019	0.032	0.017
p-value	0.716		0.085		0.851		0.102	
Visit to dentist or ophthalmologist								
	-0.002	0.004	-0.019	0.033**	-0.006	0.025	-0.002	0.009
	(0.012)	(0.012)	(0.013)	(0.013)	(0.011)	(0.017)	(0.011)	(0.016)
Baseline	0.063		0.069	0.046	0.078	0.049	0.077	0.046
p-value	0.716		0.004		0.130		0.577	
Hospitalized or surgery								
	-0.012	-0.002	-0.008	-0.003	-0.008	-0.010	-0.004	-0.017
	(0.010)	(0.009)	(0.009)	(0.013)	(0.009)	(0.016)	(0.009)	(0.015)
Baseline	0.036		0.035	0.038	0.040	0.032	0.042	0.029
p-value	0.407		0.731		0.910		0.482	
N	*53,713*		*37,592*	*16,121*	*26,858*	*26,855*	*29,529*	*24,184*

Notes: DID estimates from OLS models. Year-specific effects obtained from a model that allows the coefficient on the *Treat* × *Post* indicator to vary across the two years. Effects for remaining sub-groups obtained from split-sample estimates. Treatment and comparison groups defined as in notes to Table 1. Covariates as in notes to Table 3. 'Baseline' is the mean outcome in the sub-treatment group in the pre-reform period. p-value is for Wald test of null of equal effects across years and sub-populations. Robust standard errors adjusted for clustering at the household level in parentheses(Rogers 1994). ***, ** and * indicates significance at 1%, 5% and 10% respectively.

Table A6: Pre-reform trends in out-of-pocket expenditure

	Full sample		If use healthcare or medicines	
Treatment × 2005	0.410	0.187	1.532	0.144
	(1.304)	(0.128)	(2.502)	(0.121)
Treatment × 2006	-1.015	-0.043	-1.928	-0.113
	(1.555)	(0.132)	(2.973)	(0.124)
Model	OLS	GPML	OLS	GPML
Covariates	No	Yes	No	Yes
Sample size	*48,481*		*19,714*	

Notes: Treatment and comparison group defined as in notes to Table A1. Data from 2004, 2005 and 2006 ENAHO. Coefficients are OLS and GPML estimates of treatment group × year interactions. All models include indicators for treatment group, year 2005 and year 2006. GPML estimates additionally control for the covariates as in notes to Table 3 (except for cooking fuel, which is available only from July 2004). Robust standard errors robust to clustering at the household level in parentheses. ***, **,* indicate significantly different from zero at 1%, 5% and 10%.

Table A7: Least squares and two-part model estimates of effects on out-of-pocket (OOP) expenditure

	Full sample	If use healthcare or medicines		If OOP>0
	(1)	(2)		(3)
Including zeros				
		Correction for sample selection		
		No	Yes	
OOP	-1.267	-4.376	-4.433	
	(1.630)	(3.031)	(3.028)	
Two-part model				
P(OOP>0)	-0.022	-0.075***		
	(0.015)	(0.005)		
OOP				-2.275
				(4.547)
log OOP				-0.041
				(0.069)
Sample size	*52,812*	*22,240*		*19,548*

Notes: Top panel gives estimates from least squares regressions of the amount of OOP spending per person in 4 weeks (Soles Nuevos) including observations with no spending. The estimates in column (2) are from models without and with Heckman correction for selection into using healthcare or medicines. The selection bias correction is identified from nonlinearity. The bottom panel gives estimates of two-part model for the probability of any OOP and the positive amount. Row labelled *P(OOP>0)* gives average partial effect akin to equation (2) from logit model of probability of positive OOP. Row labelled *OOP* gives OLS estimate of impact on positive level of OOP. Row labelled *log OOL* gives estimate of semi-elasticity (exp(b)-1) obtained from OLS of log (positive) OOP. All models include covariates as in notes to Table 3. Robust standard errors corrected for clustering within households in parenthesis. ***, **,* indicate statistical significance at 1, 5 and 10%.

Table A8: Effects on conditional quantiles of out-of-pocket expenditure

Conditional quantile	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90
Full sample																	
Effect (Soles Nuevos)									0.000	-0.0658	0.0255	-0.2156	-1.2231	-2.1281	-2.3974	-3.4562	-2.3505
90% confidence interval																	
Lower bound									0.000	-0.3286	-0.5918	-1.0045	-2.2739	-3.4585	-4.5529	-5.4802	-7.8508
Upper bound									0.000	0.2890	0.4126	0.3284	0.2268	-0.3358	-0.8155	-0.5588	2.3676
Relative effect									.	-0.1805	0.0299	-0.1028	-0.2537	-0.2580	-0.1936	-0.1738	-0.0788
90% confidence interval																	
Lower bound									-1.1079	-1.0084	-0.4109	-0.3495	-0.3924	-0.3612	-0.315	-0.255	-0.2235
Upper bound									0.2182	0.3828	0.4011	0.1728	0.06115	-0.05137	-0.07353	-0.0337	0.0931
Converged b'strap iterations									482	498	499	498	499	499	499	499	499
Sample size									52,812	52,812	52,812	52,812	52,812	52,812	52,812	52,812	52,812
If use formal healthcare or medicines																	
Effect (Soles Nuevos)	-0.6218	-0.6814	-0.7051	-0.5958	-2.0107	-1.9277	-0.5866	-2.3292	-3.271	-4.2294	-4.757	-4.3221	-5.3048	-4.6962	-5.4718	-4.6879	-13.607
90% confidence interval																	
Lower bound	-0.6923	-0.7987	-1.177	-2.2037	-2.794	-3.3771	-3.5801	-4.2181	-5.1502	-5.7912	-6.3552	-6.8311	-7.9442	-9.4677	-12.337	-16.384	-22.981
Upper bound	-0.5215	-0.58	-0.2676	-0.0657	0.2705	0.5984	0.6597	-0.0083	-0.9903	-2.0347	-2.1282	-2.1266	-1.6715	-1.4516	-0.3308	1.5393	1.3841
Relative effect	-0.6324	-0.4339	-0.3161	-0.2129	-0.4019	-0.3279	-0.1027	-0.2678	-0.2958	-0.3053	-0.2868	-0.2244	-0.2181	-0.1631	-0.1533	-0.1008	-0.1829
90% confidence interval																	
Lower bound	-0.6844	-0.4868	-0.4414	-0.5045	-0.486	-0.461	-0.4192	-0.4044	-0.3994	-0.3786	-0.3503	-0.3171	-0.298	-0.2863	-0.2843	-0.2806	-0.279
Upper bound	-0.538	-0.3784	-0.1509	-0.0282	0.093	0.1641	0.1396	-0.0013	-0.1107	-0.1711	-0.1486	-0.1247	-0.0797	-0.0567	-0.0109	0.0384	0.0228
Converged b'strap iterations	499	495	485	494	493	495	498	499	499	499	499	499	499	499	499	499	499
Sample size	22,240	22,240	22,240	22,240	22,240	22,240	22,240	22,240	22,240	22,240	22,240	22,240	22,240	22,240	22,240	22,240	22,240

Notes: Censored (at zero) quantile regression estimates. Relative effect is the absolute effect as a proportion of the predicted conditional quantile under the counterfactual of there being no reform. The median proportionate effect across the treated households in the post-reform period is given for each quantile. The confidence interval is the 5 and 95 percentile estimates from 499 bootstrap samples. Bottom row gives the number of bootstrap samples for which the quantile regression converged. For computational reasons, the regressions include a simplified set of controls variables – we exclude the survey month, water source, sewage type, and main cooking fuel dummies and use a non-worker indicator instead of the four employment sector dummies, eight instead of 25 region of residence dummies, a rural area indicator instead of the eight housing density dummies, 10 instead of 12 age-gender group dummies, and three instead of six education level dummies. Like our specification to estimate effects on mean OOP spending, the quantile regressions control for seven occupation dummies, marital status, household size, and log p.c. monthly non-medical consumption. All impact estimates on mean OOP spending in Table 8 are robust to using this simplified set of control variables.