

Persistent Effects of Temporary Incentives: Evidence from a Nationwide Health Insurance Experiment

Aurélien Baillon, Joseph Capuno, Owen O'Donnell,
Carlos Antonio Tan Jr., Kim van Wilgenburg[†]

March 17, 2020

Abstract

Temporary incentives are offered in anticipation of persistent effects that are seldom estimated. We use a nationwide randomized experiment in the Philippines to estimate effects of two incentives for health insurance three years after their withdrawal. A premium subsidy had a small but highly persistent effect on enrollment. Application assistance, offered conditional on initial non-response to the subsidy, had a much larger but much less persistent effect. The subsidy persuaded those with higher initial willingness to pay to enroll and keep enrolling, while application assistance achieved a larger immediate effect by drawing in those who valued insurance less.

JEL Classification: C93, I13

Keywords: incentives, persistence, health insurance, subsidy, randomized experiment

Acknowledgements: We thank Wendy Janssens, Teresa Molina and participants of various conferences and seminars for comments. We are grateful to the European Union (FP7 program) for grant HEALTH-F2-2009-223166: Health Equity and Financial Protection in Asia (HEFPA) that funded the original experiment and to the World Bank for funding the follow-up (Daisy III) survey. The research of Aurélien Baillon is made possible by a Vidi grant of the Netherlands Organization for Scientific Research. Joseph Capuno and Owen O'Donnell are supported by the Swiss Agency for Development and Cooperation / National Science Foundation Programme for Research on Global Issues for Development through grant 400640_160374: Inclusive social protection for chronic health problems (PI: Jürgen Maurer). The experiment was registered ex-post on the AEA registry for social science experiments AEARCTR-0004977.

[†] Baillon: Erasmus School of Economics, Erasmus University Rotterdam, Tinbergen Institute baillon@ese.eur.nl; Capuno: School of Economics, University of the Philippines Diliman, jjcapuno@gmail.com; O'Donnell: Erasmus School of Economics, Erasmus School of Health Policy and Management, Erasmus University Rotterdam, Tinbergen Institute, Faculty of Economics and Business, University of Lausanne odonnell@ese.eur.nl; Tan: School of Economics, University of the Philippines Diliman, ctanjr@gmail.com; van Wilgenburg: Erasmus School of Health Policy and Management, Erasmus University Rotterdam vanwilgenburg@eshpm.eur.nl

1 Introduction

Temporary incentives for the consumption of experience goods are offered in anticipation of generating effects that persist, at least to some extent. Yet most evaluations only estimate effects achieved while the incentives are in operation. These immediate effects will not be fully sustained after the incentives are withdrawn. They may not be indicative even of relative effectiveness in the longer term if the degree of persistence varies across incentives. Smaller effects of incentives that entice only those with willingness to pay marginally below the full price may persist to a greater extent than larger effects of incentives that also draw in non-marginal types who value the good less.

This paper uses a nationwide randomized experiment in the Philippines to estimate and compare immediate and persistent effects of two temporary incentives for health insurance, a complex product that the uninitiated may have difficulty valuing prior to purchasing it. This difficulty potentially explains why take-up of insurance against health and other risks is often muted in low-income populations with little experience of insurance. In randomly selected treatment sites, households were offered a 50 percent discount on the premium of the national health insurance program for one year, plus information advertising the program's benefits and cell-phone reminders to enroll. A randomly selected half of the treatment group households that had not enrolled after nine months were offered (in addition to the subsidy) one-time assistance with application that effectively eliminated the indirect hassle costs of enrollment. We estimate immediate effects on insurance enrollment when the incentives were operating and test for the persistence of these effects three years after the incentives had been withdrawn.

Using a doubly robust estimator to correct for potential bias arising from attrition and the conditionality of the application assistance intervention, we estimate that the subsidy raised the probability of being voluntarily insured three years later by almost 5 percentage points (pp) (~100 percent relative to the counterfactual), which is more than four fifths of the immediate effect – a high degree of persistence. Application assistance had a much larger immediate effect, raising enrollment by almost 30 pp, but less than a fifth of this effect persisted. While the two incentives had very different effects in the short term, their longer-term effects differed little.

Sample respondents who were induced by the subsidy to enroll (compliers) are substantially more likely than the average respondent to have stated a high willingness to pay for insurance prior to being offered the subsidy and to have incurred medical expenses in the

previous year.¹ Immediate compliers with the application assistance intervention do not have these characteristics. This pattern is consistent with the subsidy having persuaded those who placed a higher ex ante value on insurance to enroll and to keep enrolling, while application assistance (on top of the 50 percent subsidy) achieved a larger immediate effect by inducing those with less interest in insurance, who were less likely to re-enroll when faced with the full premium and indirect costs.² The use of incentives to ameliorate adverse selection by bringing lower risks into the pool – a strategy advocated by Banerjee et al. (2019) on the basis of findings from a health insurance experiment in Indonesia – can run into a trade-off. A temporary incentive designed to reach further down into the distribution of risks may have a less sustained impact on the average insured risk. A more modest incentive can have less of an immediate effect on the risk pool but a more persistent one.

Evidence from low- and middle-income countries (LMIC) on persistent effects of temporary incentives is somewhat limited.³ Kremer and Miguel (2007) find that a temporary subsidy did not produce sustained increased use of deworming medication in Kenya. In the same country, Dupas (2014) finds persistence (after 12 months) in the effect of a temporary subsidy for another health product – insecticide-treated bed nets. The latter study, like ours, suggests that a positive learning effect from the experience of consumption dominates any negative effect from anchoring on the subsidized price (Tversky and Kahneman 1974, Simonsohn and Loewenstein 2006).⁴ We examine a complex product – health insurance – that offers even greater scope for learning through consumption. The value attached to insurance depends on comprehension of how it works, knowledge of the coverage nominally provided, experience of the reimbursement effectively delivered and observation of the quality of care that can be accessed. Little more than a tenth of our sample were familiar with the benefit package of the insurance program at baseline and less than a fifth knew the procedure for

¹ While the point estimates indicate substantial differences in the sample, these differences are not statistically significant at conventional levels.

² Sample compliers with the subsidy are substantially more likely than average to be in the top half of the living standards distribution (significant) and to live in an urban area with better access to health care – two characteristics associated with greater demand for health insurance. Compliers with application assistance are significantly less likely than average to have these two characteristics.

³ We restrict attention to evidence on persistent change in behavior after that behavior ceases to be incentivized through reduced costs of consumption. Rogers and Frey (2016) review the mixed evidence on persistent effects of behavioral interventions, mostly in high-income settings.

⁴ Dupas (2014) finds evidence of learning but not of anchoring. Spatial variation in exposure to the risk of (Malaria) infection helps identify the learning effect. There is no analogous exogenous variation in the benefit from insurance that could be used to identify the learning effect in our context. Kremer and Miguel (2007) demonstrate that anchoring is not a convincing explanation for their finding of no persistence.

making a claim. There was ample scope for learning, which could have pushed the value placed on insurance in either direction. A negative effect could have arisen, for example, from discovering that doctors charge insured patients more. Among those who enrolled in response to the offer of assistance with application, the four fifths who did not continue to insure after withdrawal of the incentives were not persuaded by the experience to raise their perceived value of the insurance sufficiently.⁵

Concern about exposure to substantial health and agricultural risks in LMIC motivates much research that seeks to understand and redress low demand for insurance.⁶ Temporary subsidization can be a remedy if knowledge and experience of how insurance operates that are acquired while consuming at the subsidized price cause a permanent outward shift in demand. Few studies test this hypothesis. Most confine attention to the immediate effect. Those that examine incentives for health insurance deliver mixed evidence (Thornton et al 2010, Das and Leino 2011, Dercon et al. 2015, Capuno et al. 2016, Wagstaff et al. 2016, Chemin 2018, Asuming et al. 2019, Fischer et al. 2018, Banerjee et al. 2019).

We know of only one other study that estimates effects on health insurance enrollment more than 18 months after the withdrawal of incentives. That study, which is based on a randomized experiment conducted in one rural district of northern Ghana, finds high persistence for three years in the effects of premium subsidies, with the smallest subsidy producing the smallest immediate effect that displayed the greatest degree of persistence (Asuming et al. 2019).⁷ Like our study, this shows that variation in persistence across interventions can change their relative effectiveness in the longer term. But given the location of the experiment in one remote district with sparse provision of medical care in one province of Ghana, it cannot be presumed that these results generalize to other settings. Indeed, two other localized health insurance experiments conducted in Nicaragua (Thornton et al. 2010)

⁵ A randomized experiment in a Kolkata slum demonstrates potential to influence demand for health insurance through experience of it (Delavallade 2017). The offer of a free preventive checkup two months before insurance was due to expire raised stated willingness to pay for insurance by 53 percent.

⁶ Platteau et al. (2017) review the evidence on the existence and causes of low demand for insurance in LMIC. For evidence of low take-up of even highly subsidized health insurance see Thornton et al (2010), Acharya et al. (2013), Banerjee et al (2014), Wagstaff et al (2016) and Chemin (2018). Capuno et al (2016) estimate immediate effects from the experiment we follow up on. Giné et al, (2008) and Cole et al, (2013) are two examples of (much) evidence of low demand for agriculture-related insurance.

⁷ One third, two thirds and full (100 percent) subsidies raised enrollment immediately by 39 percent, 48 percent and 54 percent respectively. The persistent effect is 44 percent, 30 percent and 35 percent of the respective immediate effect.

and Kenya (Chemin 2018) provide evidence of little or no persistence in the effects of incentives.⁸

The randomized experiment conducted by Banerjee et al. (2019) in two cities in Indonesia finds a more than three-fold increase in social health insurance enrollment during the year a full premium subsidy was operating. But less than a fifth of this effect persisted eight months after the subsidy had been withdrawn.⁹ Contrary to what we find in the Philippines and Assuming et al. (2018) find in Ghana, the smaller immediate effect of a partial subsidy persisted to an even lesser extent. Offering one-time assistance with application initially increased enrollment by 40 percent, but this effect did not persist at all.¹⁰

The present paper contributes evidence on persistence that is critical to establishing the effectiveness of temporary incentives by using a nationwide randomized experiment with a longer follow-up period than all but one other study of health insurance in LMIC to estimate immediate and persistent effects on enrollment in a national health insurance program. It demonstrates that the effects of temporary incentives can partially persist even three years after ceasing to operate, but the relative effectiveness of these incentives can change dramatically over time. Short-term results can be a poor guide to longer-term effectiveness not only in absolute terms but also in relative terms. Using willingness to pay elicited prior to the incentives being offered, as well as proxies for the latent demand for insurance, we show that the degree of persistence is positively associated with the ex ante value placed on insurance by compliers.

2 Health insurance in the Philippines

At the time of the health insurance experiment in 2011, the Philippines National Health Insurance Program – commonly known as *PhilHealth* – provided mandatory, contributory health insurance for formal sector salaried employees and, in principle, fully subsidized

⁸ In three street markets in Managua, free insurance for six months raised enrollment by 29 pp, but less than 10 percent of this effect persisted a year after the subsidy lapsed (Thornton et al. 2010) and the fraction fell to 5 percent after 18 months (Fitzpatrick et al. 2011). In a rural community in the Central Province of Kenya, a full subsidy (with information and application assistance) raised enrollment by 45 pp, but this effect vanished when the subsidy was withdrawn (Chemin 2018).

⁹ The modest persistence was still sufficient to raise coverage of the treatment group almost 60 percent above the low rate of coverage of the control group (6.7 percent) eight months after the subsidy had been withdrawn.

¹⁰ The immediate effect of the application assistance is less than a fifth of the effect of the full subsidy and just more than a third of the effect of the partial (50 percent) subsidy. Lower effectiveness of assisted application in Banerjee et al. (2019) is almost entirely due to failed attempts to enroll caused by deficiencies in the database.

insurance for the poor. However, responsibility for determining who was poor lay with local governments that were liable for financing part of the subsidy. As a result, coverage of the poor was patchy. The rest of the population – informal sector workers, the self-employed, the elderly and poor households that were not recognized as such by their local governments – could enroll voluntarily through the Individually Paying Program (IPP). Only one third of the eligible population took this opportunity (Manasan 2011; Capuno et al. 2016).¹¹ The purpose of the experiment was to test the effectiveness of incentives in raising this take-up rate.

In 2011, the annual premium for the IPP was 1,200 PHP (~\$30) for individuals with an average monthly income of no more than 25,000 PHP, and 2,400 PHP for those with higher incomes. Almost all paid the lower premium because it is difficult for the insurance agency to verify informal sector incomes. By 2015, the year of the follow-up survey we use to estimate persistence in the effects of the incentives, the premium had increased to 2,400 PHP for those with lower (declared) incomes and 3,600 PHP for those with higher incomes.¹² As with all PhilHealth programs, cover extends from the individual who becomes a member to their spouse, dependent children (<21 years old) and (in 2011) parents (≥ 65 years old). In the period studied, the benefit package included a wide range of inpatient services at accredited public and private hospitals, some specific outpatient treatments and limited primary care (Bredenkamp and Buisman 2016). In addition to the lack of comprehensive coverage of all treatments, the scope for health care providers to charge in excess of the reimbursement ceilings imposed by the insurer limits the effective coverage against medical expenses (Bredenkamp and Buisman 2016).

Policies implemented between 2011 and 2015 likely contributed to an increase in the population coverage of all PhilHealth programs by almost two fifths (Bredenkamp et al. 2017). This expansion of coverage should have affected our randomly generated treatment and control groups equally and so it does not jeopardize identification of the persistent effects. Local governments were relieved of their responsibility for identifying the poor and were no longer liable for co-financing the PhilHealth program that targeted this population. They were legally obliged to enroll households on a national list of the poor.¹³ Initially, this uniform, fully-

¹¹ PhilHealth claimed to cover 75% of the population through all its programs in 2010. Survey-based estimates of coverage are about 20 pp below the official figures (Bredenkamp et al 2017).

¹² The income threshold remained at 25,000 PHP.

¹³ Previously, local government units (LGU) were enjoined to enroll households on the list in PhilHealth's fully subsidized Sponsored Program. Few did so because they were partly liable for financing the subsidy.

subsidized program for the poor was financed from general tax revenue. From 2014, earmarked funding came from steep increases in tobacco and alcohol taxes. In the same year, fully subsidized insurance was extended to other disadvantaged groups (near poor, disabled, orphans etc.). At the beginning of 2015, a law came into effect that mandated government-sponsored, fully-subsidized health insurance for all senior citizens (≥ 60 years) (PHIC 2014).

3 Study design and data

3.1 Sampling and randomization

The design of the experiment is described by Capuno et al (2016), who estimate immediate effects of the two incentives. Using stratified, multistage cluster sampling, 2950 households were selected to give a nationally representative sample.¹⁴ The head of the household, or their spouse, was targeted for interview in the baseline survey, which was fielded from February to April 2011. If neither was available, then another adult (≥ 21 years) was interviewed. In each of the follow-up surveys (March–May 2012 and July–August 2015), the enumerator was instructed to interview the same person interviewed at baseline. If that person was unavailable, their spouse was to be interviewed.

Randomization of the IPP premium subsidy was done at the municipality level.¹⁵ Out of 243 randomly-selected municipalities, 179 were randomly assigned to be treatment sites and the remaining 64 designated as control sites.¹⁶ In the treatment sites, 2200 households were randomly selected for interview to establish their eligibility for the IPP and their enrollment status. Those found to be eligible but not enrolled (1037 households) were offered the insurance incentives described in the next sub-section.¹⁷ In the control sites, 730 randomly-selected

¹⁴ Stratification was by 15 regions and sub-regions. The Autonomous Region Muslim Mindanao was excluded. For each of five broad regions, and for sub-regions within them, the sample was set proportionate to the respective population. Within each (sub-)region, provinces, and then municipalities, were drawn by systematic sampling with selection probabilities determined by population sizes. Within each sampled municipality, barangays (villages/neighborhoods) were drawn by simple random sampling. A minimum of two municipalities (barangays) were drawn from each sampled province (municipality). From each sampled barangay, five households were selected by simple random sampling.

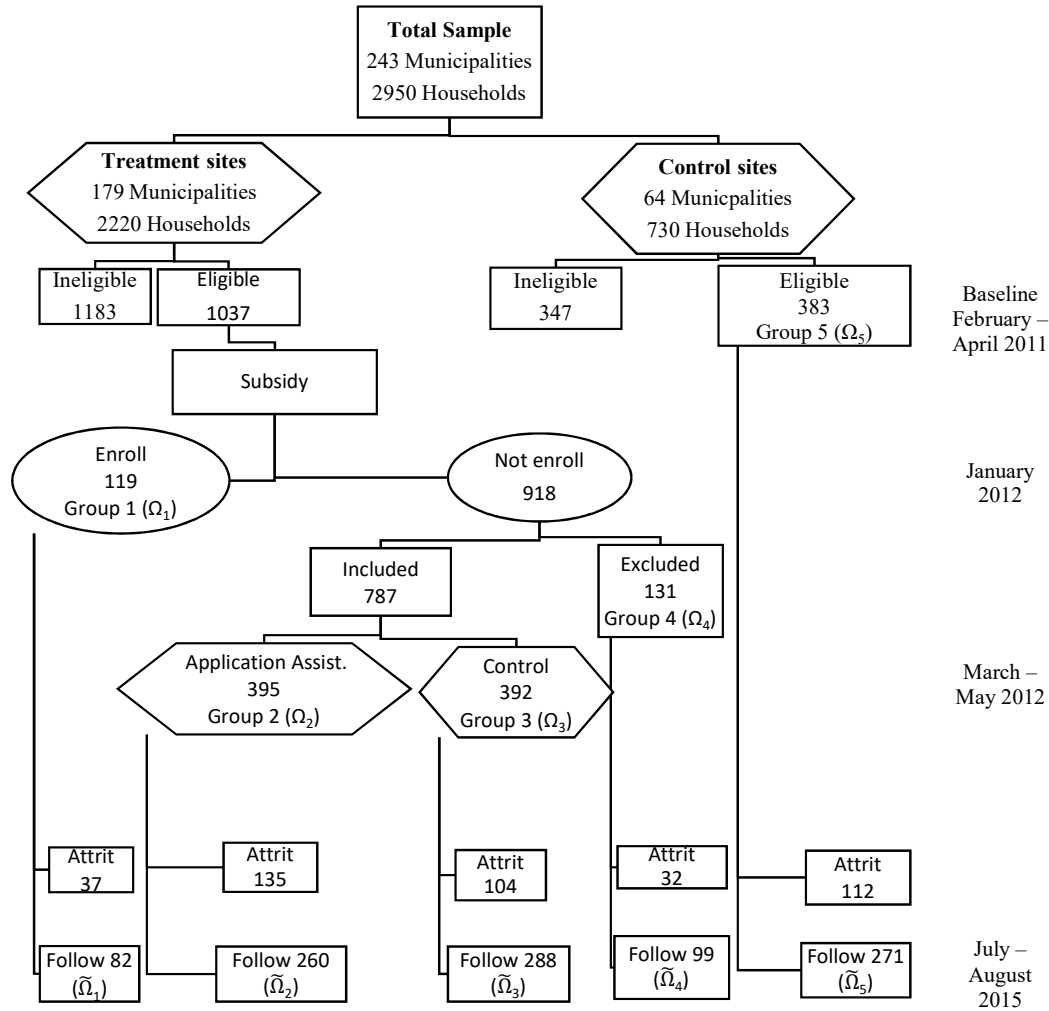
¹⁵ Municipality-level randomization reduced scope for information spillover and avoided the controls becoming disillusioned (and exiting the study) from being denied a subsidy they would have witnessed neighbors receiving.

¹⁶ The imbalance in the size of the treatment and control samples was to ensure a sufficient number of post-treatment IPP-enrolled households in the treatment sites, after allowing for ineligibility and non-compliance, to facilitate examination of how insured households coped with shocks, which was a purpose of the original study.

¹⁷ A household was deemed IPP eligible and unenrolled if neither the head nor their spouse was insured by any PhilHealth program. If either was a IPP member but had not paid the premium for six months, then coverage would have lapsed and status was set to eligible and unenrolled. If the respondent was unsure if the head/spouse were covered by any PhilHealth program, then the household was identified as eligible and unenrolled.

households were interviewed and 383 were found to be eligible but not enrolled. These control households were not offered any incentive. The participant flow is summarized in Figure 1.

Figure 1. Participant flow



Notes: “Eligible” indicates households found in the baseline survey to be eligible for the IPP insurance program but not enrolled. “Excluded” indicates refusal at baseline to take the voucher giving entitlement to enroll in the IPP at the subsidized premium. These households were subsequently excluded from the offer of application assistance. All boxes after the second row indicate numbers of households.

3.2 Interventions

At the end of the baseline interview, each of the 1037 randomly selected, IPP-eligible and unenrolled households in the treatment sites was given information on the operation of the program and offered the opportunity to enroll at a discounted premium. The household respondent was offered a voucher covering 600 PHP (\$14) of the annual premium. This was a 50 percent discount for a low income household that would qualify for the reduced premium

and a 25 percent discount for a higher income household required, in principle, to pay the full premium. In practice, since almost all who enroll in the IPP do so at the reduced premium, respondents were likely to perceive the offer as a 50 percent subsidy. The voucher was initially valid until the end of 2011 and could be used at the nearest PhilHealth office, where the application form had to be completed and the remainder of the premium paid. The voucher was not transferable.¹⁸ The respondent was also given leaflets with information about enrollment and the benefit package, as well as the application form. Until the expiry date, any recipient who had not yet redeemed their voucher was intermittently sent cell-phone messages reminding them to enroll and how to use the voucher to do so. For shorthand, we will refer to this intervention as a “subsidy”, although it also consists of information and reminders.¹⁹ Households in the control sites did not receive any information or inducements.

In January 2012, 787 households in the treatment sites that had been issued a voucher but had not yet enrolled in the IPP (according to PhilHealth’s database) were randomly allocated to one of two groups (2 and 3 in Figure 1).²⁰ Half were mailed a letter containing the same information about the program they had been given earlier and notifying them that the validity of the voucher had been extended by two months to the end of February 2012. They were also sent cell-phone messages informing them of this extension. The other half of these non-compliant households were sent a letter that, in addition to repeating the earlier information, told them that the voucher would remain valid until they were visited by a survey enumerator (March-May 2012), who would offer assistance with completion of the application form, deliver it to the PhilHealth office on their behalf and ensure that the health insurance card was mailed back to them. Essentially, this eliminated the indirect cost of enrollment, which could be substantial where transport connections were poor. We will refer to this as the “application assistance” intervention.

¹⁸ Each PhilHealth office was given a list of names of individuals in the area who had been given a voucher.

¹⁹ Banerjee et al (2019) find that two information interventions had no significant effect on social health insurance enrollment in Indonesia. This is consistent with a presumption that most of the effect of our subsidy intervention can be attributed to the premium discount. However, the type of information provided differs across the two studies. In Indonesia, all participants were informed of the benefit package, premiums and enrollment procedure, while the treatment groups were additionally given information on costs of treatments or told that insurance was mandatory and there was a waiting period after enrollment to make a claim.

²⁰ In January 2012, PhilHealth gave the study team a list of households that had used their vouchers to purchase IPP cover. Households that had received the voucher but were not on this list formed the pool from which random selection was made for the second incentive. Randomization was at the household level in this case.

At the time of the baseline survey, 131 treatment group respondents refused to accept the voucher giving entitlement to purchase the insurance at the discount price (group 4 in Figure 1). These households were excluded from the second stage of the experiment. They were not offered application assistance and are not included in the control group used to estimate its impact. They are included in the treatment group used to estimate the effect of the subsidy since they were offered this incentive.

3.4 Follow-up sample

Persistent effects are estimated using data from a second follow-up survey conducted in July-August 2015, which is more than three years after the incentives were withdrawn. The intention was to interview all 1420 households that had been IPP eligible and unenrolled at baseline. It proved possible to trace and interview 1000 of these households. The bottom row of Figure 1 shows how these households are split across the various treatment and control groups. Those lost to follow-up differ in some characteristics (see Appendix Table A1). For example, they are more likely to be urban residents, tenants and college educated. We reweight the sample to eliminate any attrition-induced compositional differences in observables between the treatment and control group households interviewed in 2015 (see section 4).

3.5 Outcomes

The main outcome of interest is whether a household is insured through enrollment of the head of household or their spouse in the IPP. Attention is restricted to households that were eligible to join the IPP but had not done so at baseline. We do not consider insurance through any other PhilHealth program because the IPP is the only one that provides coverage after voluntary payment of a premium, rather than conditional on characteristics, such as formal sector employment, old age or poverty. We are interested in whether temporary incentives to insure voluntarily had persistent effects on the decision to pay for insurance. Persistent effects are estimated using enrollment status at the time of the 2015 follow-up survey. Immediate effects are based on enrollment in 2012.

In addition to enrollment, we estimate the impact of the incentives on stated willingness to pay (WTP) for PhilHealth insurance in 2015.²¹ This outcome was elicited in two ways in the

²¹ WTP was not elicited if the respondent reported being unaware of PhilHealth. We drop these respondents when estimating effects on this outcome. Immediate effects on WTP cannot be estimated since this outcome was not elicited in the 2012 follow-up survey. It was elicited in the 2011 baseline survey and we use this information in the correction for attrition bias and in the characterization of compliers.

2015 follow-up that were randomized.²² One was an iterative bidding approach.²³ The other involved listing monetary intervals and asking the respondent to pick the one closest to their WTP.²⁴ For both methods, we assign the mid-point of the interval in which the respondent's WTP is identified to lie, unless the actual amount is reported.

4 Empirical strategy

4.1 Subsidy effect

The effects are identified principally through randomized assignment. However, we need to deal with two potential biases. One is attrition by the 2015 follow-up, which is almost 30 percent of those eligible at baseline for the IPP. The other is the conditionality of the application assistance intervention on initial non-enrollment after receiving the subsidy offer. To identify the effect of the subsidy alone, households that were also offered assistance with application must be excluded from the treatment group. This leaves three sub-groups that were exposed only to the subsidy, which are identified in Figure 1. Group 1 households responded to the subsidy by enrolling in the IPP by the end of 2011. Group 3 households received the subsidy voucher but did not enroll in the IPP by the end of 2011 and were then randomly assigned *not* to be offered application assistance. Group 4 households were offered the voucher at baseline but refused to accept it, did not enroll in the IPP and were not considered for the application assistance intervention. Even if there were no attrition, comparing the mean outcome across these three groups with the mean outcome of the control group (Group 5) would not provide an unbiased estimate of the average effect of the subsidy. Exclusion of households that were offered the subsidy, did not enroll by the end of 2011 and were subsequently randomly assigned to be offered help with application (Group 2) potentially renders the treatment group compositionally different from the control group. In the absence of any inducement, if the demand for insurance over the follow-up period by households that initially did not respond to the subsidy would have differed from the demand of the average household in the experiment, then comparing outcomes of a restricted treatment group that excludes some that did not

²² Since the elicitation method was randomized, it is orthogonal to the randomly allocated interventions. Nevertheless, to increase power, we control for an indicator of the method of elicitation.

²³ The respondent was asked whether they would pay 100 PHP per month for PhilHealth. The amount was subsequently raised or lowered and the bidding continued until the response switched. If the respondent claimed to be willing to pay more than 300 PHP, they were asked to state the amount they would be prepared to pay.

²⁴ Intervals of 50 PHP were given up to 300 PHP per month, with a final option to state the WTP amount if it exceeded that amount.

respond with outcomes of the control group will give a biased estimate. Let i index a household and define $\Omega_j = \{i \in \text{Group } j\}$ $j=1,2,3,4,5$. The full treatment group initially offered the subsidy is $\Omega_1 \cup \Omega_2 \cup \Omega_3 \cup \Omega_4$ and the control group is Ω_5 . Randomization ensures that $E[Y^0 | \Omega_1 \cup \Omega_2 \cup \Omega_3 \cup \Omega_4] = E[Y^0 | \Omega_5]$, where Y^0 is the potential outcome in the absence of any intervention, but not that $E[Y^0 | \Omega_1 \cup \Omega_3 \cup \Omega_4] = E[Y^0 | \Omega_5]$.

We deal with this potential problem by reweighting the restricted treatment group that is used to estimate the effect of the subsidy alone. Reweighting ensures that the weighted proportion of this group that enrolled by the end of 2011 is equal to the unweighted proportion of the full treatment group initially offered the subsidy that enrolled by that date.²⁵ Application of the appropriate weights to the restricted treatment group that excludes households exposed to the application assistance intervention ($\Omega_1 \cup \Omega_3 \cup \Omega_4$) ensures that households of the type that initially responded to the subsidy have the same influence on the composition of this group as these types have in the full treatment group, which is not expected to differ compositionally from the control group due to random assignment. Hence, if there were no attrition, then an unbiased estimate of the average persistent effect of the subsidy could be obtained by subtracting the mean outcome in the control group from the weighted mean outcome across the restricted treatment group that is exposed only to this intervention.

Of course, there is attrition. While it differs little, and not significantly ($p=0.5685$), between the restricted treatment group (27.0 percent) and the control group (29.2 percent) used to estimate the subsidy effect, we further reweight the sample to correct for any observable baseline differences between the subsets of these groups that are followed up. Inverse probability weights (Rosenbaum 1987, Imbens 2004) are derived from the estimated propensity scores of being offered the subsidy within the sample observed in 2015 that consists of the restricted treatment group – excluding those exposed to the application assistance intervention – plus the control group, i.e. $\tilde{\Omega}_1 \cup \tilde{\Omega}_3 \cup \tilde{\Omega}_4 \cup \tilde{\Omega}_5$, where $\tilde{\Omega}_j$ is the subset of Ω_j that contains only households observed in 2015 (see Figure 1). The propensity scores are estimated from a probit

²⁵ Define group indicators $G_{ji} = 1(i \in \Omega_j)$, where $1()$ is the indicator function, and group sizes $N_j = \sum G_{ji}$.

Appropriate weights, $w_i = G_{1i} \left(\frac{N_3 + N_4}{N_2 + N_3 + N_4} \right) + (1 - G_{1i})$, ensure $\frac{\sum_i G_{1i} w_i}{\sum_{j=1,3} \sum_i G_{ji} w_i} = \frac{N_1}{N_1 + N_2 + N_3 + N_4}$.

model of the subsidy offer on baseline values of covariates (\mathbf{X}) and the baseline value of the WTP outcome.²⁶

The average persistent effect of the subsidy can be estimated by the weighted mean difference between the (restricted) treatment group and the control group,

$$E[Y|\tilde{\Omega}_1 \cup \tilde{\Omega}_3 \cup \tilde{\Omega}_4] - E[Y|\tilde{\Omega}_5] = \frac{\sum_{j=1,3}^4 \sum_i \tilde{G}_{ji} w_i y_i}{\sum_{j=1,3}^4 \sum_i \tilde{G}_{ji} w_i} - \frac{\sum_i \tilde{G}_{5i} w_i y_i}{\sum_i \tilde{G}_{5i} w_i} \quad (1)$$

where Y is the outcome (insurance enrollment or WTP) and y_i is a realization of it observed at follow-up in 2015, $\tilde{G}_{ji} = 1(i \in \tilde{\Omega}_j)$ indicates group membership and w_i are the weights (see footnotes 25 & 26). This inverse probability weight (IPW) estimator can be obtained from a weighted least squares regression of the outcome on an indicator of the subsidy offer $\left(D_i = 1\left(\sum_{j=1,3}^4 \tilde{G}_{ji} = 1\right)\right)$ using the sample observed in 2015 that excludes those who were offered assistance with application.²⁷ We obtain our main estimates by both applying the IPW and conditioning on the baseline covariates (and WTP at baseline) in a least squares regression of the outcome on the treatment indicator. This doubly robust estimator is consistent if either the propensity score or the regression, but not necessarily both, is correctly specified (Robins and Rotnitzky 1995). We compare the main estimates with those obtained from the IPW estimator and check robustness to the exclusion of observations that lie outside the common support and those that receive very large weights. For the binary enrollment outcome, we also present probit estimates from the doubly robust estimator. For the WTP outcome, which is observed in an interval for most respondents, we also test robustness to estimating the effect by interval regression (Stewart 1983), while conditioning on covariates and applying the IPW.

²⁶ Let the propensity score for a control group household be $\Phi(\mathbf{X}_i \hat{\gamma})$, where Φ is the standard normal CDF.

That observation is given a weight $w_i = \nu_i / \bar{\nu}$, where $\nu_i = \Phi(\mathbf{X}_i \hat{\gamma}) / (1 - \Phi(\mathbf{X}_i \hat{\gamma}))$, $\bar{\nu} = \frac{1}{N_5} \sum \tilde{G}_{5i} \nu_i$, $\tilde{G}_{ji} = 1(i \in \tilde{\Omega}_j)$ and $\tilde{N}_j = \sum \tilde{G}_{ji}$, $j = 1, \dots, 5$. Weights for treated households are defined in footnote 25, and these weights are applied in estimation of the propensity scores.

²⁷ In implementing this estimator, as with all estimators used, we control for indicators of the regions used to stratify the sample.

Table 1: Balance checks in sample used to estimate subsidy effects

	Baseline mean [SD]		H ₀ : C=T p-value	Normalized difference
	Control (C)	Treatment (T)		
willingness to pay (PHP)	116.9 [58.2]	122.9 [59.0]	0.179	-0.142
total expenditure per capita (PHP)	20756 [22972]	21038 [21824]	0.870	-0.017
receive social support	0.111	0.177	0.036	-0.184
informal economic activity	0.513	0.548	0.524	-0.070
employed (head of hhold)	0.856	0.893	0.243	-0.115
college education (head of hhold)	0.077	0.077	0.971	0.003
House				
owned	0.904	0.889	0.581	0.049
# rooms	1.705	1.608	0.127	0.125
poor building materials	0.664	0.650	0.764	0.029
poor decoration	0.768	0.763	0.911	0.010
poor neighbourhood	0.638	0.580	0.333	0.119
flush toilet	0.347	0.446	0.069	-0.201
safe drinking water	0.480	0.520	0.504	-0.081
poor health (head of hhold)	0.030	0.019	0.371	0.069
adverse health event last year	0.247	0.275	0.529	-0.063
sickness / injury in past 30 days	0.269	0.264	0.904	0.011
inpatient stay in past year	0.122	0.139	0.538	-0.050
any maintenance medication	0.203	0.217	0.677	-0.036
medical expenses past 6 months (PHP)	1153 [5051]	951 [4342]	0.580	0.044
household size	5.31 [2.51]	5.29 [2.58]	0.918	0.009
# children	1.85 [2.08]	1.95 [2.16]	0.524	-0.054
>1 family in household	0.188	0.162	0.397	0.069
aware of PhilHealth insurance	0.882	0.866	0.614	0.049
aware of PhilHealth benefit package	0.118	0.154	0.229	-0.102
aware PhilHealth claims procedure	0.166	0.158	0.804	0.022
years at current address	28.13 [19.74]	27.69 [21.50]	0.776	0.025
Urban	0.376	0.493	0.161	-0.233
hospital in municipality	0.819	0.814	0.911	0.012
hospital within 1 hour	0.432	0.503	0.231	-0.143
health clinic in municipality	0.565	0.546	0.715	0.038
clinic within 15 minutes	0.199	0.264	0.089	-0.152
missing on willingness to pay	0.125	0.141	0.639	-0.045
Number of clusters	62	152	214	214
Number of households	271	469	740	740
F-test (48, 691)			1.385	(p=0.0468)

Notes: Unweighted means at baseline. Standard deviations (SD) of continuous variables in brackets. Sample used to estimate subsidy effects after attrition. Variable definitions in Appendix Table A2. Means of 15 sample stratifiers (regions) not shown. There is no significant difference in the means of any region indicator and largest magnitude of normalised difference is 0.166. F test is of joint significance of all covariates (including regions) in explaining treatment indicator. See Appendix Table B1 for weighted means, and Tables B2-B5 for means in samples used to estimate effects of application assistance and the combined treatment.

The baseline covariates used to estimate the propensity scores and as regression control variables are an extensive set of measures of household socioeconomic status, demographics, health, health care utilization and expenditure, and health insurance program knowledge, as well as location characteristics and WTP at baseline.²⁸ Table 1 shows their (unweighted) means in the restricted treatment group $(\tilde{\Omega}_1 \cup \tilde{\Omega}_3 \cup \tilde{\Omega}_4)$ and in the control group $(\tilde{\Omega}_5)$ that are used to estimate the effect of the subsidy, i.e. after dropping those offered application assistance and those lost to follow-up. Out of 48 covariates, there is only one difference in the means that is significant at the 5 percent level and only two more that are significant at 10 percent.²⁹ For all covariates, including the three for which there is a significant difference, the magnitude of the normalized difference is smaller than the 0.25 threshold often used as a rule of thumb indication of imbalance (Imbens and Rubin 2015). Despite dropping part of the treatment group and losing around 30 percent of the sample through attrition, the treatment and control groups used to estimate the effect of the subsidy appear to be reasonably balanced at baseline even before reweighting.³⁰

4.2 Application assistance effect

To estimate the persistent effect of application assistance we restrict the non-attrition sample to those who had not enrolled in the IPP by the end of 2011 despite having been offered the subsidy and compare outcomes of those randomly selected for assistance $(\tilde{\Omega}_2)$ with those who were not $(\tilde{\Omega}_3)$.³¹ Those who refused to accept the subsidy voucher $(\tilde{\Omega}_4)$ are excluded since they were not considered for application assistance. In this case, reweighting is potentially required only to correct for any attrition-induced differences in observable baseline

²⁸ In the baseline survey, WTP was elicited using the iterative list method only. For respondents who reported being unaware of PhilHealth at baseline and so were not asked their WTP for membership, we use the mean value and include an indicator of such observations.

²⁹ The F test given in Table 1 indicates that the covariates are only just jointly significant at the 5 percent level in explaining treatment.

³⁰ As would be expected, after application of the weights, there is no significant difference (at 10 percent or less) in the means for any covariate and the magnitude of the normalized difference falls to 0.06 or much less for all covariates (see Appendix B, Table B1).

³¹ There was a slight difference between the control and treatment groups in the length of the extension granted to the period of validity of the subsidy voucher. The revised expiry date was February 2012 for the controls and March – May 2012 for the treatments (depending on when they were interviewed in the first follow-up). This makes it possible that any treatment – control difference in enrollment is not entirely attributable to application assistance. However, since the subsidy had been available to both the treatment and control groups for 8-10 months when the former was offered application assistance, it seems unlikely that the estimated effect of this incentive will be biased substantially by the difference in the subsidy extension period.

characteristics. It is important to recognize, however, that the effect of application assistance in isolation can only be identified for those who did not respond to the subsidy (at least by the end of 2011). While application assistance was offered in addition to the subsidy, the control group ($\tilde{\Omega}_3$) was also offered the subsidy. We are estimating the effect of lowering the indirect costs of enrollment when the direct costs have already been reduced.

Following the general procedure described in the previous sub-section, we estimate a probit model for the probability of having been selected for application assistance conditional on being considered for this treatment and observed at follow-up, use the estimated propensity scores to construct weights for the control group and then take the weighted mean difference in outcomes. Again, we both apply the weights and condition on the baseline covariates using least squares.

There is a significant ($p < 0.1$) difference between the treatment and control groups in the means of 5 (/48) baseline characteristics (Appendix Table B2). This is around the number of differences that would be expected to occur by chance if there were no attrition.³² Attrition rates do differ: 34.2 percent (treatment) vs 26.5 percent (control) ($p = 0.03$). Despite this, the groups remain reasonably balanced – the magnitude of the normalized difference is greater than 0.25 for only 2 covariates (household per capita expenditure and household size) (Appendix Table B2).

4.3 Combined effect

We also estimate the effects of a combined treatment consisting of the subsidy offer followed, if there is no initial response, by the additional offer of assistance with application. This treatment is not simply the subsidy plus application assistance because the offer of assistance is conditional on initially not enrolling at the subsidized premium. Imposing such conditionality reduces the cost of a supplementary intervention. If there were no attrition, an estimate of the effect of this combination of incentives could be obtained by comparing the outcomes of a treatment group consisting of households that were offered the subsidy and had enrolled by the end of 2011 (Ω_1) plus those that did not respond to the subsidy and were subsequently randomly assigned to receive assistance (Ω_2) with the outcomes of a control group that was not exposed to any intervention (Ω_3). This would not provide an unbiased estimate because

³² The covariates are strongly jointly significant ($p = 0.000$) in explaining treatment. See Appendix Table B3 for the weighted means, which do not differ significantly or substantially between the treatment and control groups.

the treatment group is selected partly on response to the first incentive and can be expected to differ from the randomly selected control group. To deal with this, we again reweight the restricted treatment group to make it representative of the whole initial treatment group and so comparable with the control group.³³ Attrition-induced differences in baseline characteristics are taken into account by application of IPW and regression adjustment for covariates.

The treatment and control groups used to estimate the combined effect are again significantly different ($p < 0.1$) in the means of only 5 (/48) baseline characteristics (Appendix Table B4).³⁴ There are three normalized differences greater than 0.25 in magnitude, including that for WTP at baseline, with the treatment group stating a significantly higher value. While this indicates some imbalance in one of the outcomes at baseline, it underlines the advantage of having data available to correct for this.

4.4 Immediate effects

In addition to estimating the extent to which effects persist more than three years after the incentives were withdrawn, we estimate the immediate effects of the incentives when they were in effect. To enable comparison, we do this using the same samples (and treatment and control groups) that are used to obtain the persistent effects. Those who were lost to follow-up and not interviewed in 2015 are not used to estimate the immediate effects even if they were observed in 2012 when the immediate outcomes were measured. To estimate the immediate effect of the subsidy, we exclude those who were offered application assistance. Consequently, the two identification issues discussed in sub-section 4.1 also arise for estimation of the immediate effects and we use the same estimation methods based on reweighting. All that changes is that outcomes are measured in 2012 rather than 2015.

The respondents had initially been informed that the voucher offering the premium subsidy would expire at the end of 2011. We evaluate the immediate effect of the subsidy on insurance status in January 2012, which is before the expiry date had been extended. The immediate effect of application assistance is on insurance status in May 2012 and is estimated

³³ The weights, $\omega_i = G_{it} \left(\frac{N_2}{N_2 + N_3 + N_4} \right) + (1 - G_{it})$, ensure that the weighted share of the restricted treatment group ($\Omega_1 \cup \Omega_2$) that enrolled by the end of 2011 is equal to the unweighted share of these households in the full treatment group, $\frac{\sum_i G_{it} \omega_i}{\sum_{j=1}^2 \sum_i G_{jt} \omega_i} = \frac{N_1}{N_1 + N_2 + N_3 + N_4}$.

³⁴ The attrition rate is 33.5 percent of the treatment group and 29.2 percent of the control group ($p = 0.3177$).

through comparison of those offered this incentive with those who were not, after having restricted the sample to treatment group respondents who had not enrolled by January 2012.

All standard errors are adjusted for clustering at the municipality level, which is the level of randomization to the subsidy intervention.³⁵

5 Results

5.1 Main estimates

Both incentives succeeded, to some extent, in raising enrollment in the IPP insurance program. They did this not only while in operation but also three years after they had been withdrawn. The estimates presented in panel A of Table 2 indicate that the persistent effect of the subsidy on IPP enrollment is large relative to its immediate impact, while the much larger immediate effect of application assistance persists to a much lesser extent.

During the period that treatment group respondents could benefit from the subsidy, it raised their enrollment by 5.5 percentage points (pp), which is a 110 percent increase relative to the control group mean. The subsidy effectively offered a 50 percent price reduction and so the immediate impact on enrollment corresponds to a price elasticity of -2.2.³⁶ Two caveats should be borne in mind in interpreting this apparently substantial degree of price responsiveness. First, in addition to the premium subsidy, the intervention consisted of information and repeated reminders to enroll. Second, 89.5 percent of the treatment group chose to forgo the offer of a 50% premium discount and remain uninsured. The elasticity is large because the modest absolute increase in enrollment is from a low base. In absolute terms, a very large price reduction did not substantially reduce the uninsured rate.

The subsidy is estimated to have raised enrollment by 4.7 pp three years after it had been withdrawn and beneficiaries would have had to renew their insurance at the unsubsidized premium. This persistent effect is almost 100 percent of the control group mean and it is 85

³⁵ To be conservative, we also cluster standard errors at this level when estimating the effect of application assistance even though randomization to this treatment is at the household level.

³⁶ The subsidy corresponded to a 50 percent price reduction for those who could enroll at the reduced premium available, in principle, to low-income individuals. However, as mentioned before, effectively all informal sector workers could enroll at this premium. The arc elasticity calculated from the change in enrollment of the treatment group relative to its baseline zero enrollment and the change in price from 1200 to 600 is -3, which is calculated from $\{(5.5-0)/(5.5+0)\}/\{(600-1200)/(600+1200)\}$.

percent of the immediate effect. Apparently, most of those induced by the subsidy to enroll continued to do so after the subsidy expired.

Application assistance is estimated to have raised enrollment by 29 pp when it was offered. This is more than a six-fold increase on the control group mean and it is more than five times larger than the immediate effect of the subsidy. Clearly, offering at-home assistance with completion and submission of the application form, plus mailed receipt of the insurance card, had a very large impact on enrollment among those who were initially unresponsive to the subsidy.³⁷ This is consistent with indirect costs being a strong impediment to enrollment.

Table 2: Effects of incentives on insurance enrollment and willingness to pay

	Subsidy		Application Assistance		Combined	
	immediate	persistent	immediate	persistent	immediate	persistent
<i>A: Insured</i>	0.0550 (0.0202)	0.0469 (0.0236)	0.2912 (0.0311)	0.0536 (0.0249)	0.3613 (0.0311)	0.0954 (0.0235)
Control group mean	0.0499	0.0487	0.0426	0.0687	0.0516	0.0391
Number households	740	740	548	548	613	613
<i>B: Willingness to Pay</i>		-2.95 (6.88)		-6.97 (8.86)		-5.15 (6.60)
Control group mean		118.40		122.67		115.67
Number households		640		475		534

Notes: Panel A outcome indicates household health insurance through the PhilHealth IPP. Panel B outcome is elicited willingness to pay (WTP) per month for PhilHealth health insurance. Immediate effects are estimated using insurance enrollment in 2012 when the incentives were operating. Persistent effects are estimated using outcomes measured in 2015, more than three years after the incentives were withdrawn. Immediate effects on WTP cannot be estimated because this outcome was not measured in 2012. Sample sizes are smaller in panel B because respondents who report being unaware of PhilHealth are not asked their WTP. Estimates from doubly robust estimator that applies inverse probability weights and controls for baseline willingness to pay and covariates listed in Table 1 using weighted least squares. Control also made for sample stratification on region. Robust standard errors clustered at the municipality level in parentheses. Weights applied to obtain control group means.

After three years, those who had received the one-time offer of assistance with application continued to be more likely to insure, but the effect had fallen to less than one fifth of the immediate impact. While application assistance was much more effective than the

³⁷ Since the subsidy was offered to both the treatment and control groups used to estimate the effect of application assistance, we estimate the effect of this incentive when the premium is heavily subsidized.

subsidy when the two incentives were operating, their effects were similar in the longer term due to much greater persistence in enrollment induced by the subsidy.³⁸

The combined effect of the subsidy (plus information and reminders) followed by application assistance if the household initially did not enroll at the subsidized premium is a 36 pp increase in enrollment when the incentives were in operation. This is a seven-fold increase on the control group mean. After three years, this sequential and conditional intervention continued to have a significant positive impact on enrollment. The effect that persists is around a quarter of the immediate effect but more than double the control group mean, indicating a relatively large sustained impact on insurance.

The positive, persistent effect of the subsidy on enrollment suggests that this incentive did not backfire by anchoring willingness to pay on the subsidized price and so reducing demand when the subsidy was withdrawn (relative to what it would have been if the subsidy had never been offered). Panel B of Table 2 provides direct evidence on the effect of the subsidy (and application assistance) on WTP in 2015 that also goes against a substantial negative anchoring effect.³⁹ While the point estimate of the subsidy effect is negative, it is very small in comparison with the control group mean and not at all close to reaching significance. This does not entirely rule out a negative anchoring effect since such an effect could be offset by a positive learning effect through the experience of being insured. However, in that case, we would expect the persistent effect of the subsidy on WTP to be smaller than that of application assistance, since only the former would be affected by the negative anchoring effect. The opposite is observed.

5.2 Robustness

The results are robust to alternative methods of estimation and sample selections, as is demonstrated by the estimates presented in Table 3. The first column reproduces the main estimates from panel A of Table 2. Column (2) differs only by using probit (rather than least squares) to make the regression adjustment for the baseline covariates. Using the nonlinear estimator makes very little difference. Column (3) returns to the linear estimator but restricts the samples to common support by dropping treatment group observations with a propensity score greater than the maximum propensity score in the control group (Dehejia and Wahba

³⁸ In comparing effects of the incentives, bear in mind that the effects are estimated from different samples.

³⁹ We cannot estimate the effect on WTP in 2012 since the outcome was not elicited in the 2012 survey.

1999). Few (at most 13) observations are dropped, which is another indication that the treatment and control groups are well balanced. Dropping observations that are off the common support makes little or no difference to the estimates. Application of inverse probability weights potentially leaves estimates sensitive to control observations that attract very large weights. Column (4) tests for this by trimming the sample to exclude any control observation given a weight that is greater than one percent of the sum of all weights (Huber et al. 2013). No more than 18 observations are dropped. The estimates are a little more sensitive to this restriction, although the changes remain marginal and the general findings are robust. Column (5) gives estimates from the IPW estimator, i.e. the weighted mean difference between the treatment and control group outcomes. The estimates are very similar to those obtained from the doubly robust estimator in column (1), indicating that once adjustment is made for the covariates through application of the weights, a second adjustment through regression makes little or no difference. Column (6) gives the unadjusted treatment – control group difference in the rate of enrollment.⁴⁰ Most of the estimates are reasonably robust to this dramatic change in the estimation strategy, which again indicates the reasonable balance between the treatment and control groups. Two exceptions are that without any adjustment for covariates the estimated immediate effect of the subsidy almost doubles compared with the main estimate and the estimated persistent effect of application assistance falls by more than a third and becomes insignificant.

We conclude that there is need for covariate adjustment but the details of how it is done do not make much difference to the estimates and do not change our main findings of substantial and significant persistent effects of the incentives on insurance enrollment, with the fraction of the immediate effect that is sustained being much greater for the subsidy than for application assistance.⁴¹

⁴⁰ Adjustment is made only for the regions on which the sample was stratified to ensure that the efficiency gain from this stratification is taken into account in computation of the standard errors. This is also done for the IPW estimator.

⁴¹ The estimated persistent effects on stated willingness to pay for insurance are also highly robust to using different estimators and samples (see Appendix Table C1). The point estimates are all negative, similar in magnitude to the main estimates given in Table 2 (with the slight exception of the subsidy effect without covariate adjustment) and never close to significant.

Table 3: Effects of incentives on insurance enrollment - robustness to estimator

	Doubly robust				Inverse Probability Weights	Unadjusted
	Main Estimate	Probit	Common Support	Trimmed		
	(1)	(2)	(3)	(4)	(5)	(6)
<i>A: Subsidy</i>						
Immediate	0.055 (0.0202)	0.0617 (0.0171)	0.0585 (0.0204)	0.0465 (0.0231)	0.0582 (0.0215)	0.1053 (0.0256)
Persistent	0.0469 (0.0236)	0.0473 (0.0182)	0.048 (0.0238)	0.0577 (0.0193)	0.0503 (0.0244)	0.0506 (0.0236)
N hholds	740	740	727	729	740	740
<i>B: Application assistance</i>						
Immediate	0.2912 (0.0311)	0.2925 (0.0273)	0.2923 (0.0316)	0.2706 (0.0324)	0.2906 (0.0305)	0.3036 (0.0310)
Persistent	0.0536 (0.0249)	0.0481 (0.0227)	0.055 (0.0251)	0.0456 (0.0263)	0.0549 (0.0255)	0.0335 (0.0263)
N hholds	548	548	544	539	548	548
<i>C: Combined</i>						
Immediate	0.3613 (0.0311)	0.3611 (0.0284)	0.3633 (0.0311)	0.361 (0.0337)	0.365 (0.0343)	0.4305 (0.0341)
Persistent	0.0954 (0.0235)	0.0976 (0.0198)	0.0968 (0.0240)	0.089 (0.0253)	0.0957 (0.0245)	0.0956 (0.0284)
N hholds	613	613	605	595	613	613

Notes: Outcome is indicator of household health insurance through PhilHealth IPP. Column (1) reproduces the estimates from Table 2 obtained by applying inverse probability weights (IPW) and controlling for all baseline covariates listed in Table 1 (plus region stratifiers) in a weighted least squares regression. Column (2) is as column (1) but uses probit rather than least squares. Column (3) is as column (1) but drops treatment group observations with a propensity score greater than the maximum propensity score of the control group observations (Dehejia and Wahba 1999). Column (4) is as column (1) but drops control group observations with a weight greater than 1 percent of the sum of all weights (Huber et al. 2013). Column (5) is the weighted mean difference between the treatment and control groups without regression adjustment for covariates (other than stratification indicators). Column (6) is the unweighted mean difference between the treatment and control groups (with adjustment for stratification indicators only). All estimators control for sample stratification on (sub-)region. Robust standard errors clustered at the municipality level in parentheses.

Our estimate of the immediate effect of the subsidy on insurance enrollment is about three quarters larger than the estimate of this effect obtained by Capuno et al. (2016), who estimated only the immediate effects of the experiment interventions. The discrepancy is due to heterogeneity in the effect by attrition status, which is demonstrated in Appendix C, Table C2. Holding all other parts of the empirical strategy constant but for the exclusion of those lost to follow-up, when we impose the restriction that respondents must be observed in 2015 the estimate of the immediate effect of the subsidy increases by three quarters and the control group mean falls by 3 pp. Those lost to follow-up, who are disproportionately urban dwellers and better educated (Appendix Table A1), appear to have had greater demand for insurance in the absence of the subsidy and the incentive had less impact on this group. To estimate the

persistent effect of the subsidy, which is the goal of this paper, there is no option other than to use the non-attrition sample and comparison with the immediate effect must be made using estimates obtained from the same sample. We must accept that inference is possible only for types that are not lost to follow-up. Reweighting and further controlling for covariates through regression renders the treatment and control groups in the non-attrition sample comparable with respect to baseline characteristics, allowing unbiased estimation of the treatment effect on these types.

5.3 Heterogeneity

The subsidy had a small immediate effect on enrollment that mostly persisted, while application assistance had a much larger immediate effect that mostly failed to persist. These findings are consistent with learning from the experience of being insured having raised the perceived value of insurance to a degree sufficient to persuade immediate compliers with the subsidy to re-enroll at the unsubsidized premium but insufficient to get immediate compliers with application assistance to do so. A possible explanation for this differential persistence is that prior to becoming insured immediate compliers with the subsidy were already close to reaching the threshold willingness to pay at which they would have purchased insurance at the unsubsidized price, while immediate compliers with application assistance initially attached little value to insurance and were very far from the threshold WTP at which they would have purchased it without being incentivized. A moderately positive consumption experience would then have been sufficient for immediate compliers with the subsidy to become persistent compliers. Immediate compliers with application assistance would have required a much stronger positive learning effect to be convinced to keep purchasing when faced with the full price and non-price costs of enrollment.

To assess the plausibility of this explanation, we compare immediate compliers with the subsidy and with application assistance with respect to their stated WTP for insurance at baseline when they had not yet experienced insurance. We expect immediate compliers with the subsidy to have greater WTP than immediate compliers with application assistance.⁴² We also characterize and compare compliers with respect to baseline proxy determinants of the

⁴² Appendix D provides the logic of the expectation that the WTP interval consistent with immediate compliance with application assistance is strictly below the WTP interval necessary for immediate compliance with the subsidy. If WTP of the two groups of immediate compliers were to line up in this way, then persistent compliance with assistance after the withdrawal of this incentive would require a larger positive learning effect from being insured than would be necessary to give persistent compliance with the subsidy.

value attached to insurance, such as health indicators, past medical expenses and household resources.

Table 4 shows complier characteristics ratios (Angrist and Pischke 2009) for the two interventions.⁴³ Columns (1) and (4) give the baseline prevalence of each characteristic in the sample that is used to estimate the effect of the respective incentive. The other columns give prevalence among immediate (/persistent) compliers as a ratio of prevalence in the respective sample. For example, the top entry in column (1) indicates that at baseline prior the offer of any incentive 61 percent of the sample used to estimate the subsidy effect stated WTP of at least 1,200 PHP for PhilHealth insurance, which is the premium for those who declare low incomes.⁴⁴ The top entry of column (2) indicates that those who were induced to enroll by the subsidy were 57 percent more likely than the full sample to report WTP of at least 1,200. While the ratio is marginally short of being significantly different from 1 ($p=0.1074$), the point estimate indicates that sample respondents who complied immediately with the subsidy by becoming insured were substantially more likely than the average respondent to have reported a higher WTP for insurance. In contrast, the respective ratio for immediate compliers with application assistance is very close to 1, indicating that those who were induced to enroll by the offer of assistance were no more likely than average to have reported a high WTP. Consistent with our explanation for the differential persistence in the effects of the two incentives (see also Appendix D), the subsidy appears to have drawn a response from households that were closer to the margin of insuring even without being incentivized, while application assistance obtained its larger immediate effect by reaching further down into the distribution of preferences for insurance.

Although the ratios in the top row cells of columns (3) and (6) are not significantly different from 1, the point estimates indicate that persistent compliers in the sample with the subsidy and with application assistance are 27 percent and 19 percent, respectively, more likely

⁴³ Each ratio is the estimated effect of an incentive on enrollment in a sub-sample defined by the respective characteristic divided by the estimated effect in the full sample. Immediate and persistent complier characteristics ratios use doubly robust estimates of effects on insurance enrollment in 2012 and 2015 respectively. The respective ratios for the combined incentive are given in Appendix Table D1.

⁴⁴ We rescale reported WTP per month to annual amounts to facilitate comparison with the annual premium. At baseline, WTP was elicited using the iterative list method. A large proportion stated that they would pay 1,200 but not the next price (1,800) on the list. Still, it is somewhat anomalous that a majority of the initially uninsured sample report a WTP at least as high as the premium at which they could have insured if they had declared a low income. Some may have expected to be charged the full premium of 2,400 PHP. In any case, we are interested in the association between WTP and compliance and not in the absolute level of WTP.

than the average respondent to have reported high WTP at baseline.⁴⁵ The greater similarity across the incentives with respect to the WTP of persistent compliers is expected (see Appendix D). Unlike immediate compliance, persistent compliance is not determined by the strength of the incentive relative to pre-insurance WTP. It depends on the extent to which WTP is revised post-insurance through any learning and anchoring effects. These effects are not necessarily consistently associated with pre-insurance WTP.⁴⁶

Table 4: Complier characteristics ratios for immediate and persistent effects of incentives

Characteristic at baseline (x)	Subsidy			Application Assistance		
	Pr[x _i = 1]	Pr[x _i = 1 complier]/P[x _i = 1]		Pr[x _i = 1]	Pr[x _i = 1 complier]/P[x _i = 1]	
	(1)	immediate (2)	persistent (3)	(4)	immediate (5)	persistent (6)
A) willingness to pay ≥ 1200 PHP	0.6152 (0.0268)	1.5720 (0.3553)	1.2724 (0.4378)	0.6791 (0.0344)	1.0437 (0.0842)	1.1891 (0.3641)
B) any medical expenditure	0.2964 (0.0268)	1.6395 (0.6209)	1.2701 (0.5926)	0.3134 (0.0296)	1.0271 (0.2124)	1.7044 (0.7816)
C) ill-health	0.4969 (0.0242)	0.9930 (0.3317)	1.7429 (0.5915)	0.5056 (0.0294)	1.0648 (0.1021)	0.9328 (0.5021)
D) total household expenditure ≥ median	0.4066 (0.0258)	1.9384 (0.5379)	0.8648 (0.5458)	0.5295 (0.0356)	0.7352 (0.0994)	0.8400 (0.4883)
E) urban	0.4860 (0.0422)	1.9437 (0.5871)	1.2494 (0.4851)	0.5332 (0.0563)	0.7386 (0.1061)	0.6303 (0.5164)
Number of households	740	740	740	548	548	548

Notes: Row A) indicates willingness to pay for PhilHealth insurance of 1200 PHP or more. Row B) indicates that the household incurred medical expenses in the last six months. Row C) indicates households in which a) anyone was sick or injured in the last 30 days, OR b) there is regular monthly expenditure on maintenance medication for a chronic condition, OR c) anyone was admitted to hospital in the last year, OR d) there was any adverse health event in the last year. Row D) indicates that total household expenditure per capita above the median of the full (not analytical) sample. Row E) indicates residence in an urban location. Columns (1) and (4) give respective sample means of the characteristics. Columns (2)-(3) and (5)-(6) give the ratio of the estimated effect of the respective incentive on insurance enrollment in the sub-sample defined by the characteristic (x) to the estimated effect in the full analytical sample. Each ratio estimates prevalence of the characteristic among compliers relative to its prevalence in the full analytical sample. Estimates are obtained using the doubly robust estimator used to obtain the main estimates given in Table 2. Delta method standard errors adjusted for clustering at the municipality level in parentheses.

⁴⁵ The fact that the WTP characteristic ratio for persistent subsidy compliers is smaller (in the sample) than the respective ratio for immediate compliers implies that the subset of immediate compliers who had the most positive experience and decided to continue to insure had lower than average WTP (among immediate compliers). For application assistance, it is the opposite: persistent compliers have higher WTP than the mean of immediate compliers.

⁴⁶ The findings are maintained, and even strengthened, when examining the distribution of compliers across finer intervals of WTP (Appendix Table D2). For the subsidy, as the WTP interval is raised the complier characteristics ratios increase, although not monotonically for immediate compliers. For application assistance, the immediate complier characteristic ratio remains around 1 at all WTP intervals, while the persistent complier characteristic ratio increases monotonically with the WTP interval.

On the whole, the point estimates of complier characteristics ratios for proxy determinants of insurance demand (Table 4, rows B-E), like those for WTP, support the hypothesis that the subsidy provoked an immediate response from those who valued insurance most, while the characteristics of those induced by application assistance to enroll suggest that they had less to gain from insurance. Caution is warranted since only one of these characteristics ratios is significantly different from 1 for subsidy compliers, but the consistent pattern of the estimates lends support to the plausibility of the hypothesis.

Immediate compliers with the subsidy are 64 percent more likely than average to have incurred any medical expenses in the six months preceding the baseline survey, while immediate compliers with application assistance are no more likely than average to have done so.⁴⁷ Those who enrolled immediately in response to the subsidy are 94 percent more likely than average both to have been in the top half of the distribution of total household expenditure, which would be expected to raise the demand for insurance through an income effect, and to have been urban residents, whose proximity to more and better quality medical care would raise their demand.⁴⁸ Immediate compliers with application assistance are 26 percent *less* likely than average to have been in the top half of the expenditure distribution and to have been urban dwellers ($p < 0.05$ for both). This intervention appears to have been most effective among poorer, rural households. The one exception to the immediate response to the subsidy being greater among those anticipated to have a greater demand for (single-price) insurance is for a composite indicator of ill-health, the prevalence of which is not higher than average among immediate subsidy compliers.⁴⁹

⁴⁷ Appendix D2 shows complier characteristics ratios for three levels of previous medical expenditure (m): i) $m=0$, ii) $0 < m \leq (\text{median} \mid m > 0)$, iii) $m > (\text{median} \mid m > 0)$. For immediate compliers with the subsidy, the ratios for ii) and iii) are both greater than the ratio for i). For immediate compliers with application assistance, the ratio is smallest for iii).

⁴⁸ The indicator of total household expenditure above the median is constructed in the full sample prior to selection of the analytical samples. This explains why only 40 percent of the sample used to estimate the subsidy effect has total expenditure above the overall sample median. Finer analysis reveals that the immediate complier characteristic ratio for the subsidy is higher for the top quartile of total expenditure than it is for the second top quartile (Appendix Table D2), indicating the greatest immediate effect among the best-off households.

⁴⁹ Appendix Table D2 gives complier characteristics ratios for each component of the composite indicator of ill-health, which is defined in the notes to Table 4. This reveals that immediate compliers with the subsidy are 36 percent more likely than average to have been sick in the last 30 days and 33 percent more likely than average to spend regularly on maintenance medication. But they are less likely than average to have been admitted to hospital or to have experienced an adverse health event in the last year. Comparing immediate compliers with the subsidy and with application assistance, the characteristic ratio is larger for the subsidy for each component of the indicator separately, with the exception of experience of an adverse health event in the last year.

6 Conclusion

The persistent effects we identify would seem to imply that there is scope for using temporary incentives to permanently raise take-up of insurance against health and other risks, and possibly to increase consumption of other (undervalued) experience goods. This is an attractive policy option. Time-limited incentives are a lot cheaper than the permanent variety and they are more efficient if it is merely lack of experience that leads to sub-optimal consumption. Our findings suggest that caution be exercised, however, before reaching for this policy lever.

The potential for persistence to vary across incentives complicates the policy problem. Opting for an incentive that generates a large immediate impact by inducing even non-marginal types with low willingness to pay will be inefficient if little of the effect persists.⁵⁰ Rather than focus on the magnitude of the immediate effect, it is better to design a temporary incentive by considering the size of the potential learning effect. The incentive should compensate for the extent of undervaluation – the discrepancy between post- and pre-consumption WTP. Going beyond that will not generate marginal increases in the magnitude of the effect that persists. More estimates of persistence are needed to better inform the design of temporary incentives.

The extremely high degree of persistence we find in the effect of a temporary premium subsidy for health insurance in the Philippines is not sufficient to conclude that this policy can substantially raise insurance coverage in similar settings. The effect that persisted is large relative to the immediate effect but it is small in magnitude. A 50 percent price reduction raised enrollment immediately by only 5.5 percentage points in a sample that was initially wholly uninsured. Even though the enrollment rate remained 4.7 points higher three years after the subsidy was removed, this is hardly a substantial reduction in the high uninsured rate.

Eliminating the indirect costs of insuring – on top of the premium reduction – did raise enrollment by 30 points. But less than a fifth of this effect remained after the incentives were withdrawn. Both the size of the effect and its low persistence might be taken as evidence that application and registration costs substantially deter insurance. An obvious policy implication would call for administrative reform to facilitate enrollment and simplify re-enrollment.⁵¹

⁵⁰ Incentives that generate large immediate effects need not necessarily display low persistence. Consider a distribution of WTP that has a large mass just below the unsubsidized price. A modest subsidy can then have a large immediate effect. Even if the consumption experience causes only a modest upward revision of WTP, the effect will mostly persist. However, a larger subsidy that generates a larger immediate effect by reaching further down into the distribution of WTP will persist to a lesser extent (relative to its immediate effect).

⁵¹ Elimination of indirect costs is not necessarily optimal. The role of these costs in improving the target efficiency of subsidized social programs has long been recognized in theory (Nichols and Zeckhauser 1982) and

Certainly, there is scope for this in the health insurance programs operating in the Philippines and other LMIC. However, we are somewhat reluctant to rush to this conclusion. Another explanation for the large immediate effect of application assistance, and for the low persistence of the effect, is that respondents in the treatment group enrolled because they found it socially difficult to decline a generous offer (50 percent discount) made face-to-face by an enumerator they had invited into their home. We cannot rule out that conformity, as well as convenience, was a mechanism that helped produce the large immediate effect. This limits what we can confidently infer about the importance of non-price barriers to insurance demand. It does not, however, detract from the argument that large immediate effects generated by interventions that draw in non-marginal types who place a low value on the product cannot be expected to persist.

What caused a large fraction of those who insured due to the premium subsidy to revise their perceived value of insurance upward sufficiently to persuade them to re-enroll at the unsubsidized price? It could be that the insurance reduced exposure to the risk of incurring out-of-pocket medical expenses. It may also have made health care affordable. And the experience of being insured may have made people better informed of how insurance works. Ideally, we would test these explanations by using the randomly assigned incentives to instrument insurance and so identify its effects on medical expenditures, health care utilization and insurance knowledge. Unfortunately, the study is not powered to estimate these effects.⁵² We document persistence and explain its variation across incentives, but we cannot identify what causes it.

Our three-year follow-up on a nationwide randomized experiment reveals persistent effects of temporary incentives for health insurance. It also demonstrates that the degree of persistence can vary substantially between incentives and is likely dependent on the magnitude of the immediate effect and how close compliers are to being marginal consumers in the absence of incentives. These findings suggest that temporary incentives can potentially be effective in the longer term but only if attention is paid to how they are designed and who they target.

is increasingly demonstrated empirically (Alatas et al. 2016, Dupas et al. 2016). Even without the experiment subsidy, the health insurance we examine is offered at a reduced premium to low-income households. Given the difficulty of verifying incomes, indirect costs could potentially improve the target efficiency of the program.

⁵² The lack of power is confirmed by very imprecise IV estimates of the effects of insurance on medical expenditures, health care utilization and knowledge of insurance. No estimated effect is significantly different from zero, but effects that are large in magnitude could not be ruled out. Estimates are available on request.

References

- Acharya, A., S. Vellakkal, F. Taylor, E. Masset, A. Satija, and M. Burke. 2013. "The Impact of Health Insurance Schemes for the Informal Sector in Low- and Middle-Income Countries: A Systematic Review." *World Bank Research Observer* 28(2): 236-266.
- Alatas, Vivi, Abhijit Banerjee, Rema Hanna, Benjamin A. Olken, Ririn Purnamasari, and Matthew Wai-Poi. 2016. "Self-targeting: Evidence from a Field Experiment in Indonesia." *Journal of Political Economy* 124 (2): 371–427.
- Asuming, Patrick Opoku, Hyuncheol Bryant Kim, and Armand Sim. 2019. "Long-run Consequences of Health Insurance Promotion When Mandates are Not Enforceable: Evidence from a Field Experiment in Ghana." <https://arxiv.org/ftp/arxiv/papers/1811/1811.09004.pdf>
- Banerjee, Abhijit, Esther Duflo, and Richard Hornbeck. 2014. "Bundling Health Insurance and Microfinance in India: There Cannot Be Adverse Selection If There Is No Demand." *American Economic Review: Papers and Proceedings* 104 (5): 291-97.
- Banerjee, Abhijit, Amy Finkelstein, Rema Hanna, Benjamin A. Olken, Arianna Ornaghi, and Sudarno Sumarto. 2019. "The Challenges of Universal Health Insurance in Developing Countries: Evidence from a Large-scale Randomized Experiment in Indonesia." National Bureau of Economic Research Working Paper 26204. <http://www.nber.org/papers/w26204>
- Bredenkamp, C., T. Evans, L. Lagrada, J. Langenbrunner, S. Nachuk, and T. Palu. 2015. "Emerging Challenges in Implementing Universal Health Coverage in Asia." *Social Science & Medicine* 145: 243-248.
- Bredenkamp, C., and L.R. Buisman. 2016. "Financial Protection from Health Spending in the Philippines: Policies and Progress." *Health Policy and Planning* 31(7): 919-927.
- Bredenkamp, C., J. Capuno, A. Kraft, L. Poco, S. Quimbo, and C.A. Jr. Tan. 2017. "Expansion of Health Insurance in the Philippines: Evidence from Panel Data." World Bank, Health, Nutrition and Population Discussion Paper <https://openknowledge.worldbank.org/handle/10986/27669>
- Capuno, J.J., A.D. Kraft, S. Quimbo, C.R. Jr. Tan, and A. Wagstaff. 2016. "Effects of Price, Information, and Transactions Cost Interventions to Raise Voluntary Enrollment in a Social Health Insurance Scheme: A Randomized Experiment in the Philippines." *Health Economics* 25(6): 650–662.
- Chemin, M. 2018. "Informal Groups and Health Insurance Take-up Evidence from a Field Experiment." *World Development* 101: 54-72
- Cole, S., X. Giné, J. Tobacman, P. Topalova, R. Townsend, and J. Vickery. 2013. "Barriers to Household Risk Management: Evidence from India." *American Economic Journal: Applied Economics* 5(1): 104-35.
- Das, J., and J. Leino. 2011. "Evaluating the RSBY: Lessons from an Experimental Information Campaign." *Economic & Political Weekly* 46(32): 85-93.
- Delavallade, Clara. 2017. "Quality Health Care and Willingness to Pay for Health Insurance Retention: A Randomized Experiment in Kolkata Slums." *Health Economics* 26(5): 619–638.
- Dercon, S., J.W. Gunning, and A. Zeitin. 2015. "The Demand for Insurance Under Limited Trust: Evidence from a Field Experiment in Kenya." https://site.stanford.edu/sites/default/files/dgz_201503.pdf
- Dupas, P. 2014. "Short-run Subsidies and Long-run Adoption of New Health Products: Evidence from a Field Experiment." *Econometrica*, 82(1): 197-228.
- Dupas, P., V. Hoffmann, M. Kremer, and A.P. Zwane. 2016. "Targeting Health Subsidies through a Nonprice Mechanism: A Randomized Controlled Trial in Kenya." *Science* 353: 889.

- Fischer, Torben, Markus Frölich, and Andreas Landmann. 2018. "Adverse Selection in Low-Income Health Insurance Markets: Evidence from a RCT in Pakistan." IZA Discussion Paper No. 11751
- Fitzpatrick, A., B. Magnoni, and R.L. Thornton. 2011. "Microinsurance Utilization in Nicaragua: A Report on Effects on Children, Retention, and Health.", ILO Microinsurance Innovation Facility Research Paper No 5. Geneva: ILO.
<http://www.impactinsurance.org/sites/default/files/repaper5.pdf>
- Giné, X., R. Townsend, and J.I. Vickery. 2008. "Patterns of Rainfall Insurance Participation in Rural India." *World Bank Economic Review* 22 (3): 539–66.
- Imbens, G.W. 2004. "Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Review." *Review of Economics and Statistics* 86(1): 4-29.
- Imbens, G.W., and D.B. Rubin. 2015. *Causal Inference in Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge: Cambridge University Press.
- Kremer, Michael and Edward Miguel. 2007. "The Illusion of Sustainability." *Quarterly Journal of Economics* 122(3): 1007-1065
- Manasan, R.G. 2011. "Expanding Social Health Insurance Coverage: New Issues and Challenges." Philippine Institute for Development Studies Discussion Paper No. 2011-21.
- Nichols, Albert L. and Richard J. Zeckhauser. 1982. "Targeting Transfers through Restrictions on Recipients." *American Economic Review* 72 (2): 372–377.
- PHIC. 2014. *Implementing guidelines for the mandatory PhilHealth coverage of senior citizens pursuant to Republic Act No. 10645*. PhilHealth Circular No. 0033/2014.
http://www.philhealth.gov.ph/circulars/2014/circ33_2014.pdf, accessed 17 July 2016
- Platteau, J.-P., O. De Bock, and W. Gelade. 2017. "The Demand for Microinsurance: A Literature Review." *World Development* 94: 139-156.
- Robins, J.M., and A. Rotnitzky. 1995. "Semiparametric Efficiency in Multivariate Regression Models with Missing Data." *Journal of the American Statistical Association* 90(429): 122-129.
- Rogers, T., and E. Frey. 2016. "Changing Behavior Beyond the Here and Now." In *Blackwell Handbook of Judgment and Decision Making* (First Edition. pp. 726-748), Oxford: Blackwell.
- Rosenbaum, P.R. 1987. "Model-Based Direct Adjustment." *Journal of the American Statistical Association* 82(398): 387-394.
- Simonsohn, U., and G. Loewenstein. 2006. "Mistake# 37: The Effect of Previously Encountered Prices on Current Housing Demand." *Economic Journal*, 116(508) 175-199.
- Stewart, M. B. 1983. "On Least Squares Estimation when the Dependent Variable is Grouped." *Review of Economic Studies* 50: 737–753.
- Thornton, R.L., L.E. Hatt, E.M. Field, M. Islam, F. Solís Diaz, and M.A. González. 2010. "Social Security Health Insurance for the Informal Sector in Nicaragua: A Randomized Evaluation." *Health Economics*, 19(S1): 181-206.
- Tversky, Amos, and Daniel Kahneman. 1974. "Judgment under Uncertainty: Heuristics and Biases." *Science*, 185: 1124–1131.
- Wagstaff, A., H.T.H. Nguyen, H. Dao, and S. Bales. 2016. "Encouraging Health Insurance for the Informal Sector: A Cluster Randomized Experiment in Vietnam." *Health Economics*, 25(6): 663-674.