Results based aid for universal health coverage in poor and indigenous communities: Impact evaluation of the Salud Mesoamérica Initiative

Emily A. Dansereau

A dissertation

submitted in partial fulfillment of the

requirement for the degree of

Doctor of Philosophy

University of Washington

2019

Reading Committee:

Ali H. Mokdad, Chair

Emmanuela Gakidou

Bernardo Hernandez Prado

Tove Ryman

Maria Paola Zuniga Brenes

Program Authorized to Offer Degree:

Global Health: Metrics and Implementation Science

University of Washington

Abstract

Results based aid for universal health coverage in poor and indigenous communities: Impact evaluation of the Salud Mesoamérica Initiative

Emily A. Dansereau

Chair of the Supervisory Committee:

Ali H. Mokdad

Department of Health Metrics Sciences

Mesoamérica, a region spanning Southern Mexico to Panama, is pursuing a Universal Health Coverage (UHC) agenda to overcome significant health and economic disparities. UHC entails both ensuring financial protection from health expenditures and providing access to quality services. Beginning in 2011, eight Mesoamerican countries joined the Salud Mesoamérica Initiative (SMI), a results-based aid (RBA) program aiming to reduce maternal and child health disparities. Countries and SMI donors negotiated financially incentivized targets for improving process and outcome measures in marginalized regions. Integral to SMI is an independent evaluation, involving thousands of health facility and household surveys over several years, used for this dissertation. It first assesses baseline levels of financial protection in SMI target populations of six countries, finding high levels of out-of-pocket payments (particularly on prescription drugs) and over 20% of households forgoing care due to cost. Next, it evaluates the impact of SMI's first two phases on financial protection in four countries, finding significant improvements in Honduras (mediated by use of the public over private sector and less self-reported illness), but null or mixed results elsewhere. Finally, it evaluates SMI's impact on prenatal care in Honduras, finding that SMI improved the timing and quality of care recorded in medical records, but not according to women's own recall. Collectively, these findings give early indications that SMI RBA model improved select UHC indicators in certain settings, but not others. It implores the need for quantitative and qualitative work to evaluate if SMI can achieve population-level health improvements and understand the contextual factors leading to success in some settings and failure in others.

Aim 1: Catastrophic and impoverishing health expenditures, and care forgone due to cost, among the poorest populations: a comparative analysis of six Mesoamerican countries

ABSTRACT

Introduction: Latin America and the Caribbean is the least equitable region in the world and has the highest prevalence of catastrophic health expenditures (CHE). We aimed to compare financial protection measures for the poorest populations of Guatemala, Honduras, Nicaragua, Panama, El Salvador, and the state of Chiapas, Mexico. **Methods:** Using cross-country comparable data from 18,291 households in poor and indigenous regions, we calculated the proportion of households experiencing CHE, impoverishing health expenditures (IHE), and care forgone due to cost. We analyzed the types of expenditures driving CHE/IHE and the socioeconomic factors associated with financial protection outcomes in these vulnerable populations. **Results:** The proportion of households in our poor and indigenous sample that spent over 10% of total consumption on health ranged from 4.6% in Panama to 18.4% in El Salvador. Using the $1.90 per person per day poverty line, the proportion of households experiencing IHE ranged from 0.6% in Panama to 3.8% in El Salvador. Between 21.0% (Chiapas, Mexico) and 25.4% (Guatemala) of households had forgone care due to cost in the past year. Medication expenditure was the largest contributor to catastrophic out-of-pocket medical spending in all countries, ranging from 40.3% in Panama to 80.0% in Guatemala. Households where someone was hospitalized in the past year and households in urban areas had significantly increased odds of CHE and IHE in multiple countries. **Conclusions:** Mesoamerican countries have achieved very different levels of financial protection for their poorest citizens. Incidence of CHE, IHE, and care forgone due to cost are high even in countries with high health insurance coverage or universal free service policies. Countries must provide access to affordable medications and strengthen financial protection schemes to address CHE and IHE and ensure households do not forgo care due to cost.

INTRODUCTION

Financial protection is one of three fundamental health system objectives identified by the World Health Organization (WHO) and a Sustainable Development Goal (SDG) under the Universal Health Coverage (UHC) target. A recent analysis of 122 countries found that Latin America and the Caribbean has the highest regional prevalence of catastrophic health expenditures (CHE) [1]. It is also the world's least equitable region as measured by the Gini coefficient [2], with large pockets of poor and indigenous populations [3]. Therefore, regional discourse and strategies currently focus on achieving UHC – including financial protection – especially for marginalized groups [4].

Financial protection is a longstanding goal for many Latin American countries, several of which have undertaken significant UHC-oriented reforms over the past decades [5]. Most countries in the region now have policies affirming universal provision of care and/or comprehensive social health insurance, though their specific financing and risk-pooling structures vary widely (Table 1). For instance, the Honduran system is based almost exclusively on government provision of care, while Mexico combines public service provision with health insurance schemes including the Instituto Mexicano del Seguro Social (IMSS), Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado (ISSSTE), and Seguro Popular [6]. Robust impact evaluations have shown the success of some programs. For example, Seguro Popular reduced catastrophic expenditure by 23% overall and by 3% in poor households [7]. However, the relative performance of different countries' health system structures and programs remains an open question [8].

National-level prevalence of CHE and impoverishing health expenditure (IHE) varies widely across Latin American countries, and CHE was most recently estimated at 27.7% in Nicaragua, compared to 7.1% in Mexico and 1.4% in Guatemala [1]. CHE and IHE can also vary by income and other socioeconomic factors within a country. For example, by one definition, CHE was 12.2, 5.8, and 4.9 times more common in the poorest quintile compared to the wealthiest in Nicaragua, Mexico, and Guatemala, respectively [9]. These cross-country studies have raised legitimate concerns about the prevalence and distribution of CHE, but have the limitation of relying on country-specific survey data that were generated using unstandardized questionnaires and survey methods. This is problematic given that household health expenditure estimates are strongly influenced by the type of survey instrument used and the number of questions asked [10]. Additionally, there has yet to be a cross-country CHE study focused on poor and marginalized populations in Latin America despite high levels of inequity. Finally, most cross-country analyses have not included indirect health-related costs such as transportation, nor have they captured other indicators of financial protection including care forgone due to cost.

In this study, we compare financial protection metrics for populations living in the poorest and most indigenous regions of El Salvador, Guatemala, Honduras, Nicaragua, Panama, and Chiapas, Mexico, and examine the socioeconomic factors and specific health expenditures associated with financial protection outcomes.

METHODS

Study setting and sample

This study uses baseline household data collected for the Salud Mesoamérica Initiative (SMI) in Guatemala, Honduras, Nicaragua, Panama, El Salvador, and the State of Chiapas, Mexico [11]. SMI is a results-based financing initiative aiming to improve maternal and child health outcomes in the poorest and most marginalized regions of each country. In each country, the government and initiative partners selected municipalities with the highest proportion of households in the lowest wealth quintile and/or large indigenous populations to be part of SMI. Our research group at the Institute for Health Metrics and Evaluation served as the independent evaluator of SMI.

In our role as the independent evaluator, we conducted a baseline household survey in SMI areas. We first randomly selected census segments in SMI municipalities using probability proportional to size. We conducted a census of households in the selected segments and randomly selected eligible households (those with women 15-49 or children under 5) to participate in the survey. The sample size in each country was based on a country-specific power calculation for detecting changes in health indicators under the results-based financing scheme employed by SMI. Survey weights accounting for the multi-stage sampling design were applied for all analyses.

Data collection

The SMI baseline surveys were conducted from March 2011 to August 2013. Data were collected electronically by trained interviewers using computer-assisted personal interviewing software, and continuously uploaded to a secure database. This process allowed for ongoing systematic quality monitoring.

The household survey began with a Household Characteristics Questionnaire, which contained questions on household assets, household expenditures, medical expenditures, participation in insurance and other social programs, and languages spoken in the household (as a proxy of indigenous ethnicity). The Household Characteristics Questionnaire was conducted with the head of household or the adult with the greatest knowledge of the household as a whole; other household members could also help answer the questions. This was followed by a Maternal and Child Questionnaire, which contained questions on maternal characteristics and utilization of health services. The maternal health questions were answered by women aged 15-49, while the child health questions were answered by each child's primary caregiver. All data were collected after obtaining informed consent. The study received institutional review board approval from the University of Washington, partnering data collection agencies, and the Ministry of Health in each country.

Household expenditure data

The Household Characteristics Questionnaire requested expenditure data at several levels of granularity. At the highest level, it asked for total household expenditure. Respondents were given the option to report this for the past week, month, or year. Subsequently, it asked about expenditures for broad sub-categories, including:

- Food, alcoholic drinks, tobacco, and value of consumed food they had cultivated rather than purchased (for the past week)
- Utilities, clothing, transportation, communication, total out of pocket (OOP) medical expenditures, social insurance fees, private medical insurance fees, and other costs associated with accessing health care (for the past four weeks)
- School supplies, household maintenance and supplies, and hobbies (for the past six months)

This was followed by more granular health expenditure questions:

- Outpatient fees, inpatient fees, traditional healers, dentists, prescribed medications, un-prescribed medications, other medical supplies, lab tests, ambulance fees, other health-related expenditures (e.g., transportation, food, or childcare services necessary to obtain health care), and the value of any in-kind payments made (for the past four weeks)

The expenditure data were processed in three steps. First, all variables were converted into 2012 (the midpoint of the surveys) USD PPP and standardized to a four-week period, for instance dividing annual expenditures and multiplying weekly expenditures. Second, implausible values and extreme outliers were identified by visually inspecting the data and excluded from analysis. Third, the sum of the more granular sub-component questions were compared to the less granular reported totals. Theoretically, the granular health expenditure questions were designed to be mutually exclusive and collectively exhaustive such that they summed to the total for that category. However, this was not enforced during data collection, and the sum of the sub-components often differed from reported totals. If they could not be reconciled, we opted to use the sum of the sub-components rather than the reported total, as this is typically more complete [10]. Note that our OOP health expenditure calculations include both direct expenditures (such as fees and medications) as well as indirect expenditures (such as transportation and other expenditures associated with access).

Catastrophic Health Expenditure calculation

There are several proposed ways of calculating catastrophic expenditure [9, 12–14]. To permit comparability with recent analyses [1, 15], we used the SDG definition: OOP health expenditure less insurance reimbursements, divided by total household consumption. We used the SDG thresholds of 10% and 25% of total consumption to define CHE.

Impoverishing Health Expenditure calculation

We also calculated the prevalence of Impoverishing Health Expenditure (IHE), which measures how OOP spending relates to poverty. IHE was measured as the change in the proportion of households living below the poverty line, with and without OOP spending included in consumption. A household was considered to have faced IHE if they spent more than the poverty line based on total expenditure including OOP medical payments, but less than the poverty line based on total expenditure excluding OOP medical payments [16]. We compared results using the international extreme poverty line ($1.90 per person per day) and the poverty line commonly used for lower-middle-income countries ($3.10 per person per day). Guatemala, Honduras, Nicaragua, and El Salvador are considered lower middle income

and Mexico and Panama upper-middle-income, but a lower threshold may be appropriate as our study is among the poorest regions.

Care forgone due to cost

CHE and IHE metrics do not account for cases where medical care is so expensive, or households so poor, that care is forgone entirely. To capture this, our survey asked: "For the past 12 months: Did any household member did not seek medical care due to cost?" We report on this indicator and examine relationships between CHE, IHE, and care forgone due to cost.

We also examined the relationship between CHE, IHE, and two proxy measures of overall service utilization (whether a sick woman or sick child received care for their illness in the two weeks prior to survey) to examine the influence of non-financial barriers to care on CHE and IHE.

Factors associated with CHE, IHE, and care forgone due to cost

We examined the univariate and multivariable associations between socioeconomic characteristics and CHE, IHE, and care forgone due to cost in our study populations. Socioeconomic characteristics were selected based on previous associations identified in the literature and *a priori* hypotheses developed by researchers on our team. Independent variables included rural vs. urban location; head of household gender; presence of infant household member (<1 year old); presence of elderly household member (>60 years old); household size; indigenous ethnicity (based on language spoken by head of household); maternal education (maximum achieved by a woman in the household); wealth quintile (measured by an asset-based wealth index created through principal component analysis); household member hospitalized in the past year; time to reach usual health facility; and health insurance status. Participation in the Oportunidades conditional cash transfer program (now known as Prospera) was included for Chiapas. We first examined the bivariate relationships between each socioeconomic variable and outcome. Next, we examined the same covariates (regardless of if there was a significant bivariate relationship) in multivariable logistic regression models (the outcomes again being CHE, IHE and care forgone due to cost). Finally, we added binary indicators for whether the household reported certain types of expenditures (outpatient, inpatient, medications, traditional medicine, other direct medical expenditures, and transportation/access expenses). Models were country-specific, and standard errors were adjusted to account for the survey design. Households missing data for one or more covariates were excluded from all estimates and analyses reported in this paper. Analyses were conducted using Stata 13.1 [17].

RESULTS

Sample description

We analyzed data from 18,291 households in the poor and indigenous regions of Guatemala, Honduras, Nicaragua, Panama, El Salvador, and the State of Chiapas, Mexico (Table 2). The majority of households were in rural areas, ranging from 100% rural households in Panama to 60% in Chiapas, Mexico. Most heads of household (HoH) spoke an indigenous language (either exclusively or in addition to Spanish) in

Panama (97%), Guatemala (76%), and Chiapas (67%), whereas all Honduran households and 82% of those in Nicaragua spoke only Spanish. Between 40% and 50% of households had an infant under 1 year old in all countries except El Salvador (21%). Between 10% and 20% of households had a member over 60 years old, with the exception of Panama (43%).

Most respondents traveled <15 minutes to their usual health facility in Panama (65%). In contrast, more than half (51%) of Nicaraguan households traveled over 30 minutes to their usual facility, including 28% who traveled more than an hour. Nicaraguan households were most likely to have had a member hospitalized in the past year (18%), followed by Honduras (13%), Chiapas (8%), Panama (8%), and Guatemala (5%) (this question was not asked in El Salvador). 64% of Chiapan households were enrolled in Oportunidades. As expected, Chiapas had the highest insurance coverage (89%), with 83% of households insured through Seguro Popular and 6% through other schemes such as IMSS, ISSSTE, and private insurers. Far fewer households were insured in Panama (18%), Nicaragua (15%), Guatemala (15%), El Salvador (9%), and Honduras (2%).

Prevalence of CHE, IHE, and care forgone due to cost

The proportion of households in our sample that experienced a CHE at the 10% threshold was 4.6% in Panama, 8.9% in Nicaragua, 9.7% in Honduras, 9.9% in Guatemala, 13.8% in Chiapas, and 18.4% in El Salvador (Table 3). The proportion of households experiencing CHE was lower using the 25% threshold, though the relative rank of countries was the same (2.5% in Panama to 8.9% in El Salvador). The rank ordering of countries was also the same for IHE, whether the poverty line was set at $1.90 (0.6% in Panama to 3.8% in El Salvador) or $3.10 (1.2% in Panama to 3.3% in El Salvador).

The proportion of households where a member had forgone care due to cost in the past year was 21.0% in Chiapas, 21.1% in Honduras, 21.9% in Panama, 24.6% in Nicaragua, and 25.4% in Guatemala. This information was not captured in the El Salvador survey. Taking CHE, IHE, and care forgone due to cost together, the proportion of households experiencing any type of financial protection failure ranged from 24.0% in Panama to 30.8% in Guatemala.

At the household level, there was overlap between experiencing CHE and forgoing care due to cost (Supplemental Figure S1). Households experiencing CHE (at either the 10% or 25% level) were more likely to have forgone care due to cost than those without a CHE (47% among 10% CHE households vs. 23% non-CHE households in Guatemala, 30% vs. 20% in Honduras, 32% vs. 19% in Chiapas, 41% vs. 23% in Nicaragua, 55% vs. 20% in Panama). Likewise, households forgoing care due to cost were more likely to have experienced a CHE (18% of household that forwent care due to cost vs. 7% of those that did not forgo due to cost in Guatemala, 14% vs. 9% in Honduras, 21% vs. 12% in Chiapas, 15% vs. 7% in Nicaragua, 12% vs. 3% in Panama).

There was also overlap between CHE/IHE and other proxy indicators of service utilization behaviors in our sample of poor and indigenous households. In Guatemala, Honduras, Chiapas, and Nicaragua, children who received care for a recent fever or cough were more likely to be from a household that experienced a CHE in the past year than children who were sick and did not receive care (Supplemental Figures S2a-d). For example, among Chiapan households where a child had fever or cough in the prior

two weeks, 25% of children that received care were from a household that experienced 10% CHE in the past year, compared to 11% of children that did not receive care for the illness. This relationship was inconsistent for recent women's illnesses (Supplemental Figures S3a-d). In Guatemala, Honduras, and Panama, women who received care for their illness were more likely to be from a household experiencing CHE than women who did not have their illness treated. The opposite was true in Chiapas and Nicaragua. No differences were significant. The majority of households forgoing care for children or women reported a non-cost-related barrier; the most common barrier was lack of medication availability.

Health expenditure patterns of CHE and non-CHE households

In all countries, most of the poor households in our study did not make any OOP medical payments (73.0% in El Salvador, 79.2% in Chiapas, Mexico, 81.0% in Nicaragua, 83.5% in Honduras, 84.2% in Guatemala, 88.8% in Panama). Of the households that did have some OOP medical expenditure, the majority were considered catastrophic at the 10% level. In all countries, less than 10% of all households had made a non-catastrophic OOP payment.

As expected, among households with any OOP payment, the median amount of these payments was substantially higher among households that experienced a CHE than those that had not. For example, in Guatemala, the median non-catastrophic OOP expenditure was $3.45, compared to $20.67 for the median catastrophic OOP expenditure.

Medication: On average, medication costs were the largest contributor to OOP medical expenditures for households experiencing CHE at the 10% threshold (Figure 1). Specifically, prescribed drugs constituted the majority of medication expenditures in every country, as opposed to over-the-counter drugs. While this held true in all countries, total medications expenses were a relatively smaller contributor in Panama (40.3% of health expenditures among 10% CHE households, on average), and a relatively larger contributor in Guatemala (80.0%). Medications (particularly prescribed medications) were also the largest contributor when using the 25% CHE threshold or IHE indicators (Supplemental Figures S4a-c). Most households experiencing a CHE reported some expenditure on either a prescribed or non-prescribed medication (from 76.0% in Panama to 92.8% in Guatemala), while most non-CHE households did not (4.1% of all non-CHE households in Panama to 10.3% in Nicaragua). The median amount of these total medication expenditures was also substantially higher for CHE households in most countries (e.g., $4.44 non-CHE vs. $18.50 CHE in Nicaragua), with the exception of Panama ($1.70 non-CHE vs. $2.75 CHE).

Outpatient and inpatient fees: Outpatient and inpatient fees each comprised <7% of OOP payments by 10% CHE households in all countries, with the exception of outpatient fees in El Salvador (10.2%) and Honduras (9.8%) (Figure 1). For households meeting the 25% threshold, outpatient fees comprised over 15% of OOP payments in Honduras. While relatively rare, inpatient and outpatient fees have the potential to be extremely high. While only 5.5% of Panamanian households experiencing CHE reported any outpatient fees, the median amount was $44.70 for those who did. Similarly, only 8.2% of Guatemalan households experiencing a CHE had any inpatient fees, but the median amount was $59.08.

It is also important to note that this category only includes service fees, and not the other costs associated with outpatient or inpatient care such as medications or tests.

Traditional medicine: Panama was the only country where traditional medicine (as self-defined by the respondent) contributed substantially to CHE in our sample (30.6% of health expenditures in CHE households) (Figure 1). 43.1% of Panamanian households experiencing CHE reported some type of traditional medicine expense, compared to only 2.2% of households that did not experience CHE.

Access: Access and transportation costs comprised <5% of health expenditures in CHE households in Guatemala, Honduras, Chiapas, and Nicaragua, but was a relatively important factor for CHE households in Panama (11.2% of health expenditures) and El Salvador (14.8% of health expenditures) (Figure 1). However, access spending patterns were different for these two countries: El Salvadorian households experiencing a CHE were twice as likely to report spending money on access compared to those in Panama (41.0% and 21.7%, respectively), but the median amount of these expenditures was nearly three times as high in Panama ($16.09) as in El Salvador ($5.61).

Other medical expenditures: Other medical expenditures comprised over 13% of OOP costs for households experiencing CHE in the study regions in Chiapas, Nicaragua, and El Salvador. The composition of this "other" group varied by country. In Chiapas and El Salvador, the largest "other" proportion was for labs and diagnostics; in Nicaragua, it was for dentistry.

Household characteristics associated with CHE, IHE, and care forgone due to cost

Households where a member had been hospitalized in the past year had significantly greater odds of experiencing CHE at the 10% threshold for the study populations in all countries (except El Salvador, where this question was not asked) (Table 4). In Honduras, Nicaragua, Panama, and El Salvador, households in the wealthiest and/or second-wealthiest quintile (with quintiles defined among the poor and indigenous households in the study area) had increased odds of CHE as compared to the poorest quintile. Households where a woman had a post-primary education had increased odds in Guatemala (OR: 2.53, 95% CI: 1.67–3.81). Households with health insurance had increased odds of CHE in Guatemala (OR: 2.30, 95% CI: 1.72–3.07). Indigenous-language-speaking households had increased odds of CHE in Panama compared to those that also spoke Spanish (OR: 3.55, 95% CI: 1.18–10.69); in Guatemala, households speaking both Spanish and an indigenous language had increased odds compared to households speaking only Spanish (OR: 2.24, 95% CI: 1.47–3.42). Results were of similar direction and magnitude when the outcome was CHE at the 25% threshold, with the addition that having an elderly household member was associated with increased odds of 25% CHE in Honduras and El Salvador (Supplemental Table S1). Hospitalization in the past year was also a significant predictor of IHE at the $1.90 and $3.10 level in most countries; however, wealth quintile was not a significant predictor for the IHE outcomes (Supplemental Tables S2 and S3).

Including additional covariates for whether the household had made any health expenditures in different categories confirmed that having a medication expenditure (of any amount) was significantly associated with greater odds of CHE in all countries, as was making any outpatient payments (Supplemental Table S4). Making inpatient payments was significantly associated with greater odds of

CHE in all countries but Honduras, and making access payment was significantly associated with greater odds of CHE in Honduras, Chiapas, and El Salvador. Traditional medicine payments were only significantly associated in Panama.

Having a household member hospitalized in the past year was associated with increased odds of forgoing care due to cost in all countries (Table 5). In all countries but Nicaragua, households in the wealthiest quintile (relative to the other poor/indigenous households) had significantly lower odds of having forgone care due to cost compared to the poorest quintile. This was also true of households where a woman had post-primary education in Panama. Having more household members was associated with higher odds of forgoing care due to cost in Guatemala (OR: 1.05, 95% CI: 1.02–1.08), Honduras (OR: 1.10, 95% CI: 1.05–1.16), Nicaragua (OR: 1.05, 95% CI: 1.00–1.09), and Panama (OR: 1.06, 95% CI: 1.01–1.11). Households more than an hour from their usual health facility had higher odds in Chiapas, Mexico (OR: 1.43, 95% CI: 1.02–1.99). In Guatemala, odds were also greater for households that spoke only an indigenous language (OR: 1.80, 95% CI: 1.25–2.60) and those with health insurance (OR: 2.01, 95% CI: 1.51–2.67).

DISCUSSION

This is first study to provide cross-country comparable measures of financial protection for the poorest and indigenous populations of Mesoamerica. The incidence of CHE and IHE in these populations varied widely by country; at the extremes, CHE and IHE were approximately four times more common in El Salvador compared to Panama. Medication expenditures and hospitalization are shared drivers of CHE across the SMI areas of each country. Insurance programs do not appear to be providing sufficient financial protection, and over a fifth of households in our study recently forwent care due to cost. Several salient points arise from our findings.

First, it is clear that OOP medical payments severely and disproportionately affect the poor and indigenous populations in our study. In most countries, we found much higher CHE and IHE incidence compared to national averages calculated with an equivalent methodology (Supplemental Table S5) [18]. For example, 9.9% of households we interviewed in Guatemala spent over 10% of their total consumption on health, compared to only 1.4% in a 2014 national survey. The same was true in Chiapas (13.8% in our study populations, 7.13% in a 2012 national survey) and Panama (4.6% in our study, 1.4% in 2008 national study). Comparisons with national surveys from Nicaragua present an unclear situation. A 2014 survey found an extremely high national CHE incidence of 27.7%, compared to only 8.9% in our study. However, this is contradicted by another analysis that found Nicaraguan households in the poorest quintile were 12 times as likely to experience CHE compared to the richest quintile [9]. It is difficult to reconcile these variable findings, so it is possible that recent surveys in Nicaragua should not be compared to our survey due to methodological differences.

Second, we find that medication expenditures, and specifically prescription medications, were the top contributor to OOP for these poor populations in all countries. This is cause for concern, given that every country in our study has policies to provide essential medications at low or no cost to the poorest

populations. This disconnect may be attributed to frequent medication stock-outs; we documented these stock-outs through health facility surveys that were conducted over the same areas and time period as the household data used for this study. For example, at baseline, only 54% of health facilities in the areas we surveyed in Chiapas had ampicillin or amoxicillin in stock [19]. Improving supply chain reliability in poor and remote regions was a major focus in the first phase of SMI, and an interim survey conducted 18 months after baseline found major improvements in supply availability. For example, the availability of ampicillin and amoxicillin rose to 98% in Chiapas. It will be important to examine whether improvements in medication supply have resulted in reduced incidence of CHE and IHE for these populations.

Third, insurance programs in the region are not achieving financial protection for the poorest populations. Chiapas, Mexico, is the only place in our study to have widespread insurance coverage, including the Seguro Popular program specifically targeted to the poorest populations. While past analyses have shown that Seguro Popular reduces OOP and catastrophic expenditure [20], we find that CHE and IHE incidence are higher among Chiapas' poorest populations than among the poorest populations in many other countries. One explanation is that the effect of Seguro Popular differs depending on a household's location, access to care, and types of expenditures [21, 22]. In particular, there is evidence Seguro Popular does not reduce OOP expenditure due specifically to medication [23], which appears to be the main driver of cost for this population. This may also relate to the stock-out of medications: if options provided by Seguro Popular are not available, households must turn to other sources that are not covered by the scheme. In the other countries, insurance coverage is extremely low, and in fact is associated with *increased* odds of CHE in Guatemala. This somewhat paradoxical finding indicates that adverse selection is at play and these poor households only enroll in insurance programs when they face severe medical costs. A more comprehensive risk-pooling approach would be needed for a sustainable insurance program.

Finally, CHE and IHE must be interpreted in the context of utilization. The Sustainable Development Goals conceptualize financial protection and health service coverage as two complementary components of UHC. The relationship between utilization and expenditure is complex. In the context of CHE and IHE, we find it important to measure care forgone due specifically to cost, as this constitutes a failure in financial protection. Choosing to forgo care is an indication that the household viewed the potential cost as catastrophic. There are a number of potential relationships between CHE/IHE and forgone care. First, we find evidence that CHE/IHE is influencing utilization behaviors: when examining the question of care forgone due to cost over the past year, we find evidence that households experiencing CHE/IHE are more likely to forgo care due to cost, suggesting that CHE/IHE can lead to lower utilization. Second, we find evidence of a relationship in the opposite direction, where utilization behaviors are influencing rates of CHE/IHE: many households are forgoing care due to cost or other reasons without experiencing a CHE/IHE. This indicates that if all poor households received the care they needed, the incidence of CHE and IHE would likely increase substantially, particularly among households with a cost-related barrier. While we cannot know whether the forgone costs would have been catastrophic or impoverishing by our standard definitions, the fact that the household chose not to make the expenditures seems a good indication that the household themselves may have seen it as potentially catastrophic.

The relationship between utilization and expenditure can also help explain why households in the wealthiest quintile (relative to other households in the poor study population) have increased odds of CHE (but not IHE) in most countries and reduced odds of forgoing care due to cost in some countries: experiencing a CHE requires having enough funds available to make any type of payment at all. Therefore, we suggest that financial protection monitoring should account for households that are forgoing care due to cost. In particular, while Nicaragua had the second lowest CHE incidence in our study, it also had one of the highest rates of care forgone due to cost; the reasons for this are unclear, given a national policy of universal free health care. Beyond recent hospitalization and having an extremely low wealth index, our models found few significant predictors of forgoing care due to cost, indicating that the drivers of this behavior are poorly understood.

More broadly, non-financial barriers to care are also an important factor affecting financial protection metrics. It is logical that low utilization will drive down CHE rates, and this is evidenced in our household-level data using the proxy of utilization for recent child and women's illness (though with the limitation that these proxies were over only a two-week period and we did not capture information on illness severity). The utilization-CHE dynamic also plays out at the country level: for example, while Panama had the lowest CHE and IHE incidence in our study, Panama also had very low coverage of services such as vaccination, antenatal care, post-natal care, and modern contraception in the study population as compared to other countries in the study [24]. As interventions (including SMI) work to increase access to and utilization of services in the most vulnerable populations, it is important to concurrently monitor whether this also increases the incidence of IHE and CHE.

Our findings should be interpreted in light of several limitations. First, our CHE and IHE definitions rely on arbitrary thresholds, and there is ongoing debate as to how well these indicators measure the effects of health expenditures on a household's well-being. We have chosen to align with the SDG definitions – and it is encouraging that countries rank similarly regardless of the threshold used – but more longitudinal research on the effects of health expenditures on household well-being is needed. In terms of representativeness, it is important to note that the sample was designed to represent households with women of reproductive age and/or young children, in the poorest regions of each country. Due to the way these regions were selected, the study may tend to capture more of the rural than the urban poor. These considerations limit comparability with data on the poorest quintile that are often reported by nationally representative surveys, but do have practical advantages in terms of informing geographically targeted interventions and align with national priorities around maternal and child health. Finally, our expenditure data are based on recall and are therefore subject to measurement error and bias. However, we have minimized differential bias across countries through the use of consistent questionnaires, enumerator training activities, and survey protocols across countries.

CONCLUSIONS

Our study indicates that Mesoamerican countries have not achieved universal health coverage for their poorest populations. In the poorest areas, nearly any care that is not free meets the standard definition of being catastrophic for the household. The most common OOP health expenditures are on prescription

and non-prescription medications, but systems must also account for access costs such as transportation. Furthermore, over a fifth of households surveyed in each country had forgone care due to the cost in the past year, including in settings with prominent insurance programs (e.g., Chiapas) and universal free service policies (e.g., Nicaragua). Substantial changes are needed to ensure financial protection and access to care for the poorest populations.

ACKNOWLEDGMENTS

This study was funded by the Bill & Melinda Gates Foundation, the Spanish Agency for International Development Cooperation, and the Carlos Slim Health Institute, through the Inter-American Development Bank. We thank all the children and families who willingly participated in the study. We thank central and local governments for the support they extended to the study teams and their facilitation of access to communities and health facilities. We thank all the agencies, teams, and supervisors who conducted data collection, including El Colegio de la Frontera Sur-Mexico, The Centre for Research and Health Studies at the National Autonomous University of Nicaragua (CIES-UNAN), UNIMER, and Fundación FES.

FIGURES

Figure 1: Median amount and average composition of OOP health expenditures, among households in the poor and indigenous SMI regions with any OOP health expenditures.

TABLES

Table 1: National health financing and financial protection indicators

Table 2: Characteristics of sampled households in the poor and indigenous SMI study regions

Table 3: Proportion of households in the poor and indigenous SMI regions experiencing CHE, IHE, and care forgone due to cost, by country

Table 4: Household factors associated with spending over 10% of total expenditure on health in the poor and indigenous SMI regions

Table 5: Household factors associated with having forgone care due to cost in the past year in the poor and indigenous SMI regions

CITATIONS

1. Wagstaff A, Flores G, Hsu J, Smitz M-F, Chepynoga K, Buisman LR, et al. Progress on catastrophic health spending in 133 countries: a retrospective observational study. Lancet Glob Health. 2018;6:e169–79.

2. World Bank. GINI index (World Bank estimate). 2017. https://data.worldbank.org/indicator/SI.POV.GINI. Accessed 4 Oct 2017.

3. Romero Reynó J, Castelló González M, Choonara I. Child health in Central America and the Caribbean. Arch Dis Child. 2015;100 Suppl 1:S70–1.

4. PAHO. Strategy for Universal Access to Health and Universal Health Coverage. 2014. http://www.paho.org/uhexchange/index.php/en/uhexchange-documents/technical-information/26-strategy-for-universal-access-to-health-and-universal-health-coverage/file.

5. Atun R, de Andrade LOM, Almeida G, Cotlear D, Dmytraczenko T, Frenz P, et al. Health-system reform and universal health coverage in Latin America. Lancet Lond Engl. 2015;385:1230–47.

6. Baeza C, Packard T. Beyond Survival: Protecting Households from Health Shocks in Latin America. Palo Alto, CA/Washington, DC: Stanford University Press/World Bank; 2000.

7. King G, Gakidou E, Imai K, Lakin J, Moore R, Nall C, et al. Public Policy for the Poor? A Randomised Assessment of the Mexican Universal Health Insurance Programme. The Lancet. 2009;373:1447–54.

8. Laurell AC. Universal health coverage in Latin America? Soc Med. 2016;9:54–5.

9. Knaul FM, Wong R, Arreola-Ornelas H, Méndez O, Network on Health Financing and Social Protection in Latin America and the Caribbean (LANET). Household catastrophic health expenditures: a comparative analysis of twelve Latin American and Caribbean Countries. Salud Publica Mex. 2011;53 Suppl 2:s85-95.

10. Lavado RF, Brooks BP, Hanlon M. Estimating health expenditure shares from household surveys. Bull World Health Organ. 2013;91:519-524C.

11. Mokdad AH, Colson KE, Zúñiga-Brenes P, Ríos-Zertuche D, Palmisano EB, Alfaro-Porras E, et al. Salud Mesoamérica 2015 Initiative: design, implementation, and baseline findings. Popul Health Metr. 2015;$metaTagArticleVolume:3.

12. Xu K, Evans DB, Kawabata K, Zeramdini R, Klavus J, Murray CJL. Household catastrophic health expenditure: a multicountry analysis. Lancet Lond Engl. 2003;362:111–7.

13. Wagstaff A, van Doorslaer E. Catastrophe and impoverishment in paying for health care: with applications to Vietnam 1993-1998. Health Econ. 2003;12:921–34.

14. O'Donnell O, van Doorslaer E, Wagstaff A, Magnus L. Analyzing Health Equity Using Household Survey Data. In: Analyzing Health Equity Using Household Survey Data. Washington, DC: World Bank; 2008. p. Chapter 18. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.430.3297&rep=rep1&type=pdf.

15. Wagstaff A, Flores G, Smitz M-F, Hsu J, Chepynoga K, Eozenou P. Progress on impoverishing health spending in 122 countries: a retrospective observational study. Lancet Glob Health. 2018;6:e180–92.

16. WHO. The Global Reference List of 100 Core Health Indicators: Headcount ratio of impoverishing health expenditure. 2015.
http://www.who.int/healthinfo/indicators/2015/chi_2015_124_health_expend_impoverishing.pdf.

17. StataCorp. Stata Statistical Software: Release 13. College Station, TX: StataCorp LP; 2013.

18. WHO Global Health Observatory. Catastrophic out-of-pocket health spending (SDG indicator 3.8.2) - Data by country. 2017.
http://apps.who.int/gho/data/view.main.UHCFINANCIALPROTECTION01v?lang=en. Accessed 24 Mar 2018.

19. Mokdad AH, Palmisano EB, Zúñiga-Brenes P, Ríos-Zertuche D, Johanns CK, Schaefer A, et al. Supply-side interventions to improve health: Findings from the Salud Mesoamérica Initiative. PloS One. 2018;13:e0195292.

20. Galárraga O, Sosa-Rubí SG, Salinas-Rodríguez A, Sesma-Vázquez S. Health insurance for the poor: impact on catastrophic and out-of-pocket health expenditures in Mexico. Eur J Health Econ. 2010;11:437–47.

21. Sosa-Rubí SG, Salinas-Rodríguez A, Galárraga O. [Impact of "Seguro Popular" on catastrophic and out-of-pocket health expenditures in rural and urban Mexico, 2005-2008]. Salud Publica Mex. 2011;53 Suppl 4:425–35.

22. Grogger J, Arnold T, León AS, Ome A. Heterogeneity in the effect of public health insurance on catastrophic out-of-pocket health expenditures: the case of Mexico. Health Policy Plan. 2015;30:593–9.

23. Wirtz VJ, Santa-Ana-Tellez Y, Servan-Mori E, Avila-Burgos L. Heterogeneous Effects of Health Insurance on Out-of-Pocket Expenditure on Medicines in Mexico. Value Health. 2012;15:593–603.

24. Mokdad AH, Gagnier MC, Colson KE, Zúñiga-Brenes P, Ríos-Zertuche D, Haakenstad A, et al. Health and wealth in Mesoamerica: findings from Salud Mesomérica 2015. BMC Med. 2015;13:164.

Figure 1

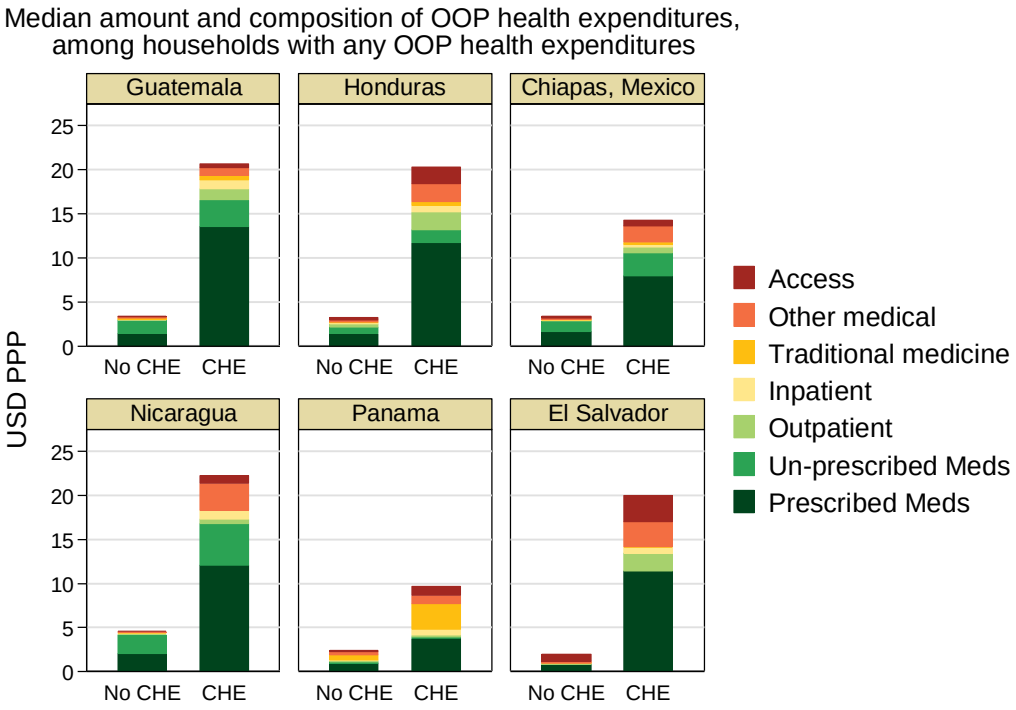

Median amount and composition of OOP health expenditures,
among households with any OOP health expenditures
Guatemala
Honduras
Chiapas, Mexico
Nicaragua
Panama
El Salvador
USD PPP
0
5
10
15
20
25
No CHE
CHE
Access
Other medical
Traditional medicine
Inpatient
Outpatient
Un-prescribed Meds
Prescribed Meds

Table 1: National health financing and financial protection indicators

Country	Health expenditure breakdown			% of government expenditure on health	Health expenditure per capita	Health insurance as % of current health expenditure		Insurance coverage in paper's study population (SMI areas)
	OOP	Public	Private			Social/ compulsory	Voluntary	
Guatemala	53%	38%	10%	18%	$473	16%	4%	14%
Honduras	44%	51%	6%	15%	$400	12%	5%	2%
Mexico	44%	52%	4%	12%	$1,122	28%	5%	89%*
Nicaragua	38%	56%	6%	24%	$445	24%	2%	15%
Panama	22%	73%	5%	15%	$1,677	25%	7%	16%
El Salvador	29%	66%	5%	17%	$565	29%	6%	9%

Sources: NHA Indicators; World Bank Open Data

*In Mexico, SMI areas are only within the state of Chiapas

Table 2: Characteristics of sampled households in the poor and indigenous SMI study regions

	Guatemala	Honduras	Chiapas, Mexico	Nicaragua	Panama	El Salvador
N	3,878	2,661	5,027	1,980	1,279	3,466
Urban	16%	16%	40%	36%	0%	26%
HoH language						
Spanish only	24%	100%	33%	82%	3%	NC
Spanish and indigenous	69%	0%	59%	18%	71%	NC
Indigenous only	6%	0%	7%	0%	26%	NC
Household size (median)	6	5	5	5	7	5
Female HoH	14%	21%	10%	27%	30%	30%
Any infant <1	50%	43%	45%	43%	49%	21%
Any elderly 60+	15%	15%	10%	13%	43%	18%
Travel time to usual health facility						
<15 min	31%	32%	33%	23%	65%	21%
15-29 min	28%	25%	29%	26%	7%	28%
30-59 min	22%	20%	23%	22%	9%	29%
60+ min	19%	22%	15%	28%	19%	22%
HH member hospitalized in past year	5%	13%	8%	18%	8%	NC
Has health insurance	15%	2%	89%	15%	18%	9%
Enrolled in opportunidades	NA	NA	64%	NA	NA	NA
Maternal education (HH max)						
None	28%	7%	13%	9%	11%	8%
Primary	51%	66%	49%	43%	44%	50%
Post-primary	21%	27%	38%	48%	46%	42%
Household expenditures						
Total monthly household expenditure per capita (median USD PPP)	$48	$58	$41	$66	$59	$53
% HH with any direct medical expense	16%	16%	20%	19%	11%	23%
Direct medical expense amount (median USD PPP))	$10	$9	$9	$8	$3	$13
% HH with any indirect medical expense (transit, access)	1%	3%	2%	1%	1%	12%
Indirect medical expense amount (median USD PPP)	$7	$9	$5	$8	$6	$3
% HH with any lost wages due to illness	5%	5%	7%	5%	1%	NC
Lost wages due to illness (median USD PPP)	$10	$15	$9	$11	$15	NC

NC: Not Collected; NA: Not Applicable

Table 3: Proportion of households in the poor and indigenous SMI regions experiencing CHE, IHE, and care forgone due to cost, by country

	CHE: 10%	CHE: 25%	IHE: $1.90	IHE: $3.10	Care forgone due to cost	Any financial protection failure
	% of households (standard error)					
Guatemala	9.9% (0.5%)	4.7% (0.4%)	2.3% (0.3%)	1.9% (0.2%)	25.4% (0.7%)	30.8% (0.8%)
Honduras	9.7% (0.6%)	4.3% (0.4%)	1.8% (0.3%)	1.9% (0.3%)	21.1% (0.8%)	28.0% (0.9%)
Chiapas, Mexico	13.8% (0.5%)	6.0% (0.4%)	2.7% (0.2%)	2.1% (0.2%)	21.0% (0.6%)	30.7% (0.7%)
Nicaragua	8.9% (0.7%)	3.5% (0.4%)	1.9% (0.3%)	1.5% (0.3%)	24.6% (1.1%)	30.1% (1.1%)
Panama	4.6% (0.9%)	2.5% (0.7%)	0.6% (0.2%)	1.2% (0.4%)	21.9% (1.7%)	24.0% (0.2%)
El Salvador	18.4% (0.7%)	8.9% (0.5%)	3.8% (0.3%)	3.3% (0.3%)	NC	NC

Table 4: Household factors associated with spending over 10% of total expenditure on health in the poor and indigenous SMI regions

	Guatemala		Honduras		Chiapas, Mexico		Nicaragua		Panama		El Salvador	
Urban	1.37	[0.82,2.28]	1.04	[0.62,1.76]	1.19	[0.90,1.59]	0.53*	[0.31,0.91]	NA		1.14	[0.82,1.57]
HoH language[a]												
Spanish only	Ref		NA		Ref		Ref				NA	
Spanish and indigenous	2.24***	[1.47,3.42]	NA		0.86	[0.68,1.08]	0.61	[0.24,1.54]	Ref		NA	
Indigenous only	0.69	[0.27,1.76]	NA		0.76	[0.50,1.13]			3.55*	[1.18,10.69]	NA	
Household size	1.01	[0.95,1.07]	1.00	[0.94,1.06]	1.01	[0.96,1.05]	0.98	[0.91,1.06]	1.10	[0.93,1.29]	0.96*	[0.93,0.99]
Female HoH	1.17	[0.84,1.63]	0.97	[0.67,1.41]	0.89	[0.65,1.23]	1.12	[0.67,1.88]	1.19	[0.63,2.23]	0.94	[0.77,1.16]
Any infant <1	1.09	[0.84,1.43]	0.86	[0.62,1.19]	1.03	[0.87,1.23]	1.31	[0.89,1.94]	0.92	[0.63,1.35]	1.15	[0.92,1.43]
Any elderly 60+	1.07	[0.85,1.35]	1.13	[0.79,1.63]	1.17	[0.89,1.53]	0.86	[0.51,1.47]	0.73	[0.28,1.92]	1.15	[0.92,1.42]
Travel time to usual health facility												
<15 min	Ref		Ref		Ref		Ref		Ref		Ref	
15-29 min	0.86	[0.57,1.30]	0.94	[0.65,1.35]	0.89	[0.67,1.17]	1.19	[0.66,2.16]	1.46	[0.51,4.17]	1.16	[0.90,1.49]
30-59 min	0.83	[0.55,1.25]	0.97	[0.59,1.60]	1.06	[0.81,1.39]	0.96	[0.53,1.76]	0.63	[0.17,2.35]	1.12	[0.84,1.49]
60+ min	0.92	[0.62,1.38]	1.02	[0.70,1.48]	1.14	[0.84,1.53]	1.17	[0.65,2.13]	1.15	[0.51,2.60]	1.29	[0.90,1.86]
HH member hospitalized in past year	5.11***	[3.13,8.34]	2.28***	[1.57,3.31]	2.23***	[1.67,2.98]	2.14***	[1.56,2.95]	2.81**	[1.49,5.31]	NC	
Has health insurance	2.30***	[1.72,3.07]	NA		1.06	[0.79,1.41]	1.12	[0.70,1.79]	0.66	[0.31,1.43]	NA	
Enrolled in opportunidades	NA		NA		1.04	[0.84,1.28]	NA		NA		NA	
Maternal education												
None	Ref		Ref		Ref		Ref		Ref		Ref	
Primary	1.30	[0.92,1.85]	1.69	[0.77,3.71]	1.09	[0.83,1.43]	0.86	[0.33,2.28]	0.95	[0.37,2.46]	0.87	[0.60,1.26]
Post-Primary	2.53***	[1.67,3.81]	1.88	[0.85,4.18]	1.19	[0.86,1.65]	1.31	[0.52,3.32]	0.46	[0.15,1.45]	1.12	[0.77,1.64]
Wealth quintile												
1 (Poorest)	Ref		Ref		Ref		Ref		Ref		Ref	
2	0.74	[0.47,1.15]	0.98	[0.60,1.58]	1.13	[0.81,1.57]	1.39	[0.74,2.63]	6.37	[0.98,41.48]	0.83	[0.60,1.15]
3	0.79	[0.50,1.25]	1.17	[0.73,1.87]	1.15	[0.81,1.63]	1.87	[0.98,3.57]	3.03*	[1.03,8.92]	1.01	[0.72,1.44]
4	0.88	[0.55,1.42]	1.11	[0.62,2.01]	1.24	[0.88,1.74]	2.02*	[1.05,3.88]	5.21*	[1.01,26.75]	1.35	[1.00,1.83]
5 (Wealthiest)	0.92	[0.56,1.51]	1.68*	[1.02,2.76]	1.41	[0.97,2.06]	3.76***	[1.86,7.61]	3.55	[0.70,18.06]	1.68**	[1.21,2.35]
N	3878		2661		5027		1980		1279		3466	

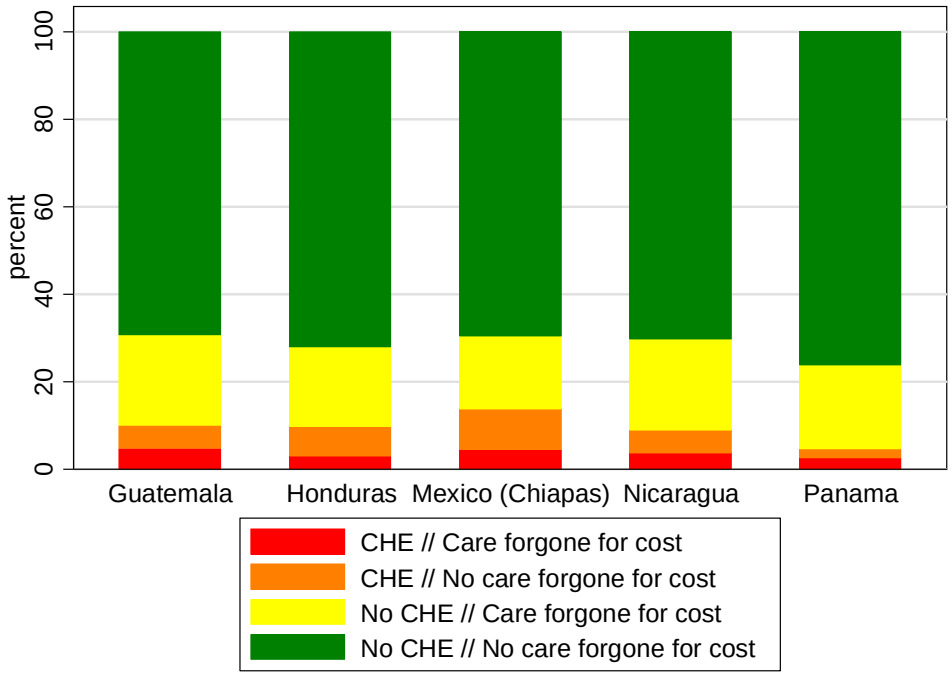
NA: Not applicable, NC: Not collected. [a]In Panama, no households that speak only Spanish (3% of households) had a CHE; we combined Spanish-only speaking with the Spanish and indigenous-speaking category in the regression analysis and compared with the indigenous-only category.

Table 5: Household factors associated with having forgone care due to cost in the past year in the poor and indigenous SMI regions

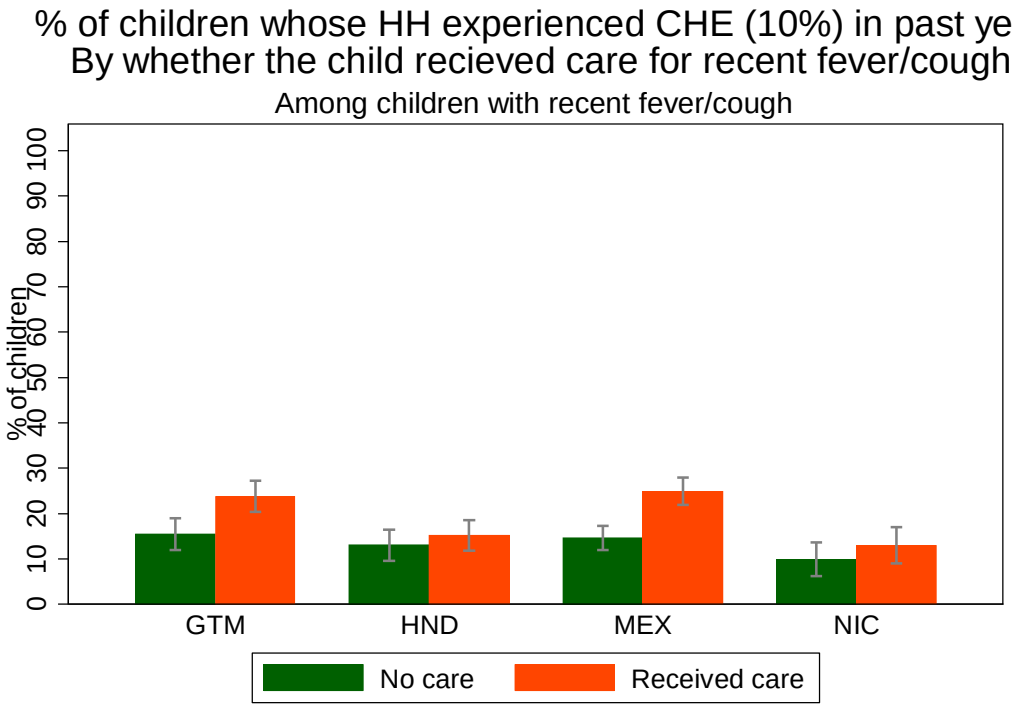
	Guatemala		Honduras		Chiapas, Mexico		Nicaragua		Panama	
Urban	1.40	[0.98,1.99]	1.33	[0.99,1.80]	1.16	[0.88,1.53]	0.85	[0.57,1.27]	NA	
HoH language[a]										
Spanish only	Ref		NA		Ref		Ref			
Spanish and indigenous	1.20	[0.93,1.55]			0.96	[0.75,1.22]	0.92	[0.63,1.35]	Ref	
Indigenous only	1.80**	[1.25,2.60]			0.75	[0.53,1.07]			1.16	[0.80,1.69]
Household size	1.05***	[1.02,1.08]	1.10***	[1.05,1.16]	1.02	[0.99,1.05]	1.05*	[1.00,1.09]	1.06*	[1.01,1.11]
Female HoH	1.19	[0.92,1.53]	0.94	[0.72,1.23]	1.17	[0.93,1.47]	1.16	[0.86,1.56]	1.19	[0.88,1.60]
Any infant <1	0.94	[0.79,1.12]	0.81*	[0.67,0.98]	0.96	[0.82,1.13]	0.94	[0.75,1.16]	1.25	[0.88,1.79]
Any elderly 60+	1.09	[0.90,1.31]	1.21	[0.93,1.59]	1.11	[0.84,1.46]	1.30	[0.97,1.76]	0.73	[0.48,1.12]
Travel time to usual health facility										
<15 min	Ref		Ref		Ref		Ref		Ref	
15-29 min	1.10	[0.87,1.40]	1.02	[0.80,1.30]	0.99	[0.79,1.24]	1.10	[0.82,1.47]	1.15	[0.49,2.69]
30-59 min	1.06	[0.81,1.40]	1.06	[0.74,1.52]	1.00	[0.78,1.30]	0.92	[0.67,1.27]	0.91	[0.36,2.28]
60+ min	1.03	[0.76,1.39]	1.17	[0.82,1.67]	1.43*	[1.02,1.99]	1.02	[0.72,1.44]	1.88	[0.87,4.04]
HH member hospitalized in past year	2.54***	[1.78,3.62]	1.58**	[1.19,2.08]	1.74***	[1.37,2.20]	1.60***	[1.23,2.09]	2.07*	[1.11,3.86]
Has health insurance	2.01***	[1.51,2.67]	1.00	[1.00,1.00]	0.87	[0.69,1.11]	0.76	[0.48,1.22]	1.33	[0.69,2.54]
Enrolled in opportunidades					1.01	[0.84,1.22]				
Maternal education (highest in HH)										
None	Ref		Ref		Ref		Ref		Ref	
Primary	1.07	[0.90,1.27]	0.77	[0.54,1.09]	1.03	[0.83,1.26]	1.26	[0.82,1.92]	1.09	[0.68,1.74]
Post-Primary	0.91	[0.69,1.20]	0.72	[0.49,1.05]	1.05	[0.80,1.36]	1.21	[0.81,1.83]	0.57*	[0.36,0.89]
Wealth quintile										
1 (Poorest)	Ref		Ref		Ref		Ref		Ref	
2	0.94	[0.75,1.18]	1.21	[0.90,1.62]	0.84	[0.65,1.09]	0.82	[0.59,1.14]	2.63	[0.95,7.28]
3	0.79	[0.61,1.03]	0.86	[0.58,1.29]	0.95	[0.72,1.24]	0.89	[0.60,1.32]	0.76	[0.25,2.31]
4	0.86	[0.65,1.13]	0.69	[0.46,1.04]	0.85	[0.64,1.14]	0.93	[0.62,1.41]	1.02	[0.32,3.19]
5 (Wealthiest)	0.70*	[0.52,0.93]	0.51***	[0.35,0.75]	0.65*	[0.46,0.91]	0.71	[0.42,1.19]	0.47*	[0.23,0.96]
N	3878		2661		5027		1980		1279	

NA: Not applicable, NC: Not collected. [a]In Panama, no households that speak only Spanish (3% of households) had a CHE; we combined Spanish-only speaking with the Spanish and indigenous-speaking category in the regression analysis and compared with the indigenous-only category. The question about utilization was not asked in El Salvador.

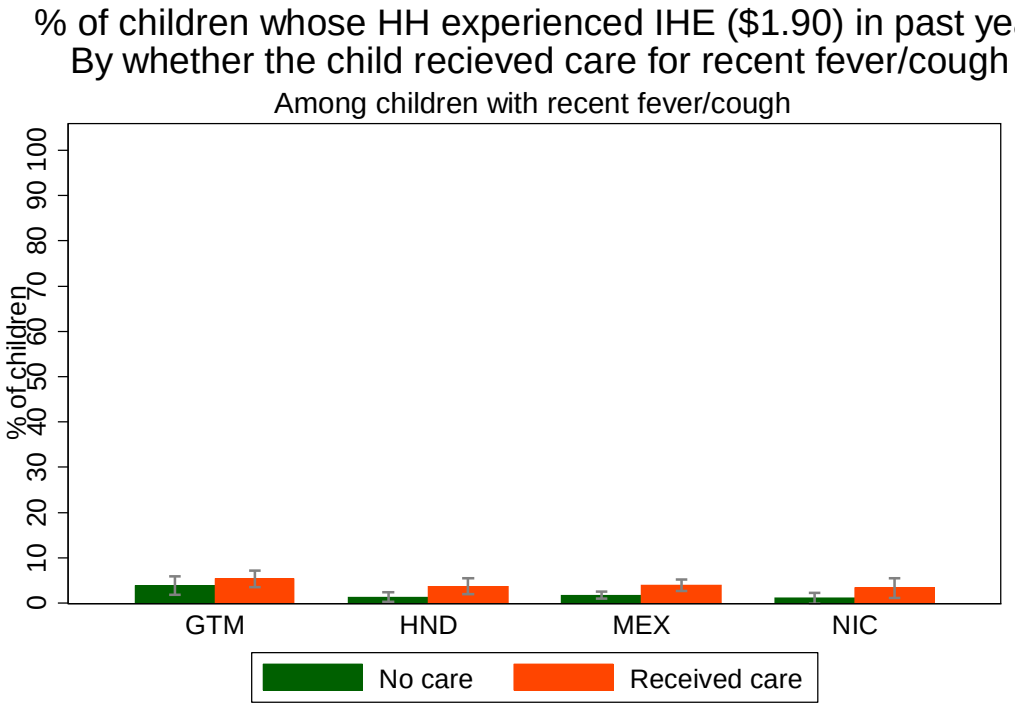
Supplemental Figure S1: Overlap of CHE (10% threshold) and care forgone due to cost at the household level in SMI study populations, by country



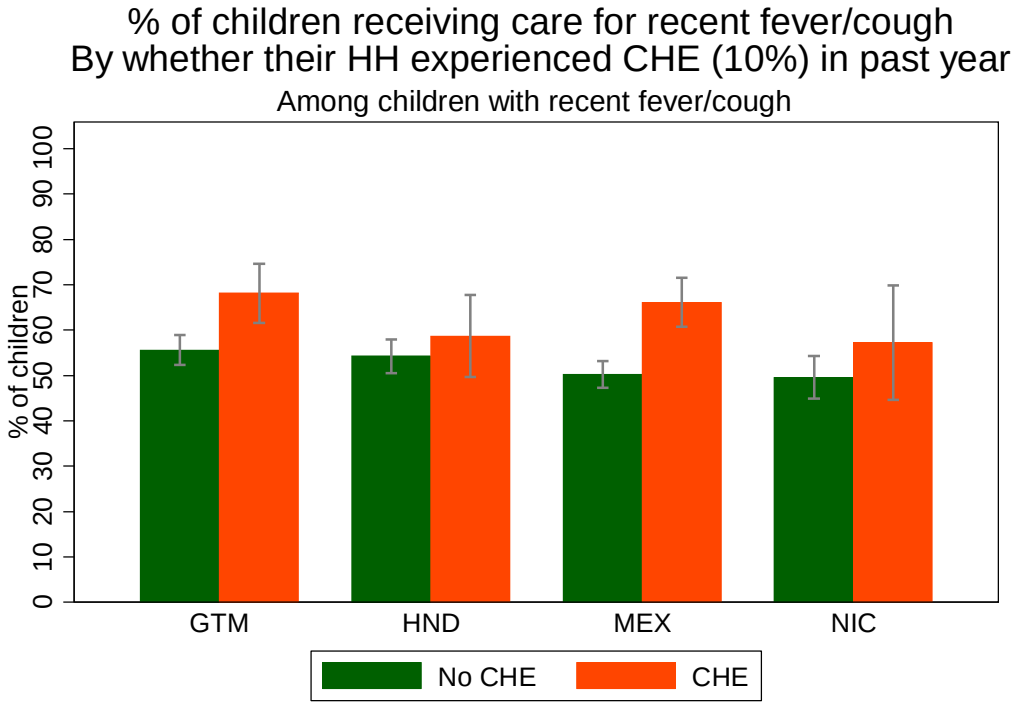
Supplemental Figure S2a



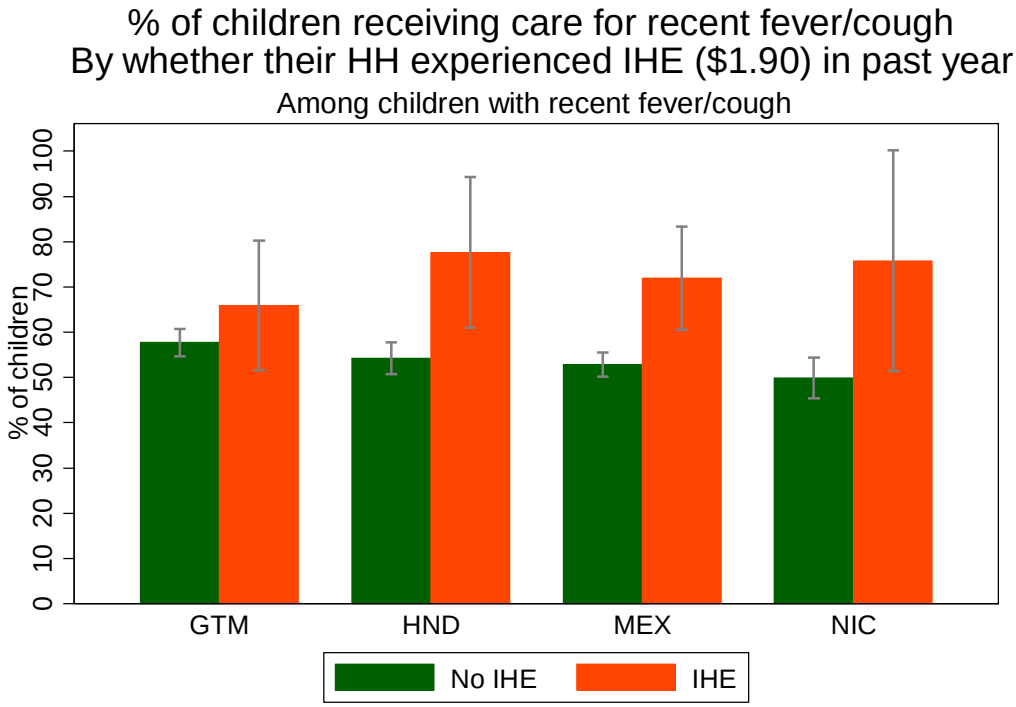
Supplemental Figure S2b



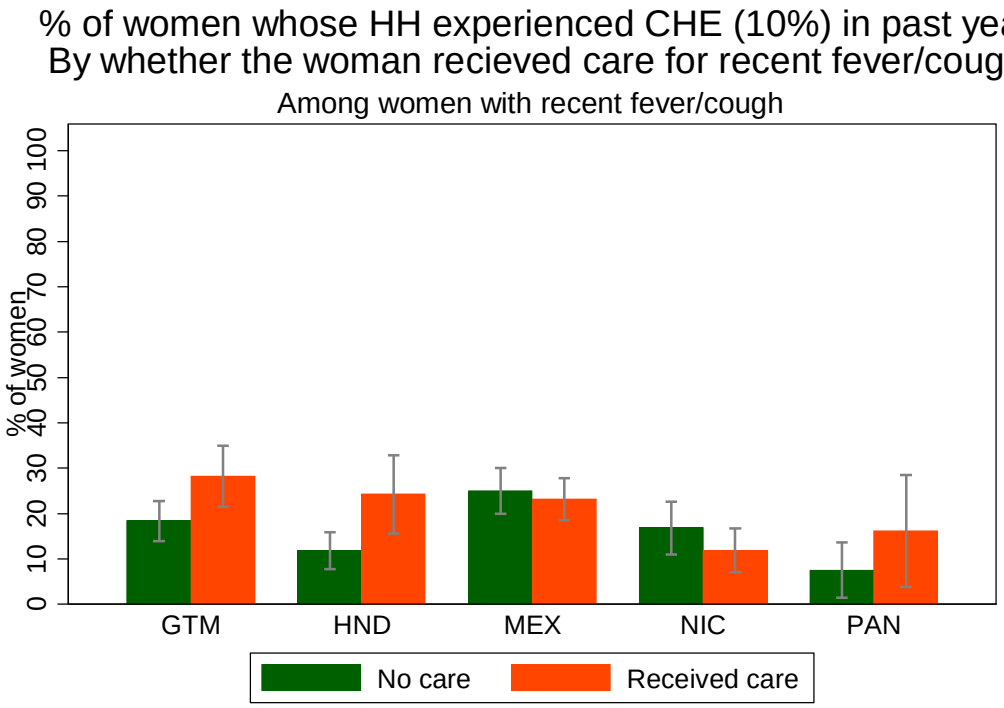
Supplemental S2c



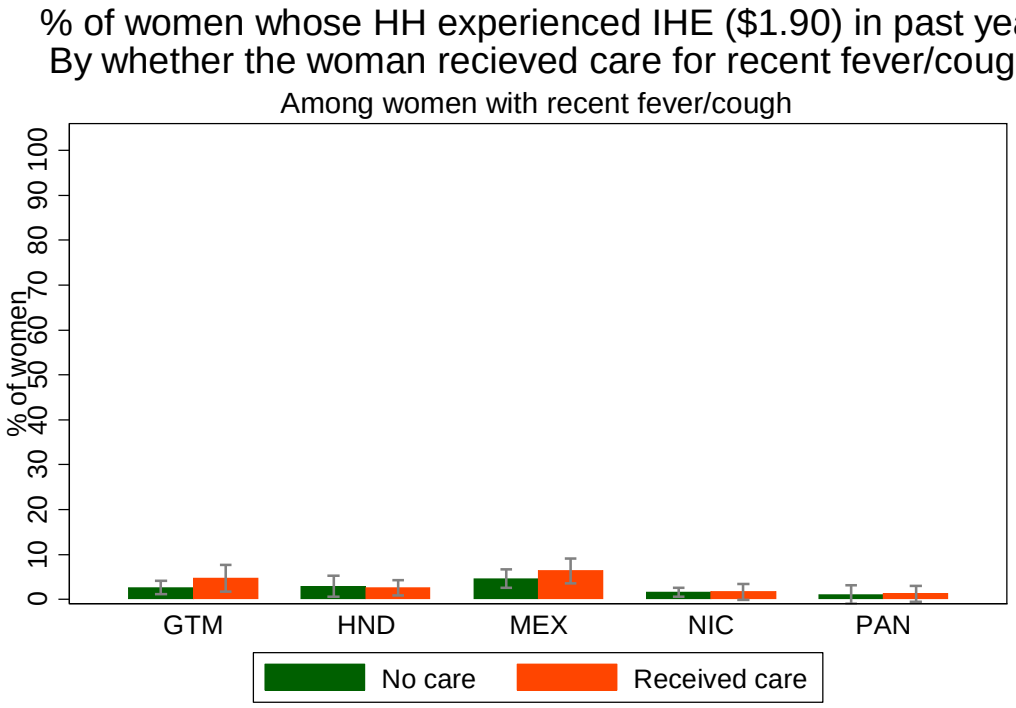
Supplemental Figure S2d



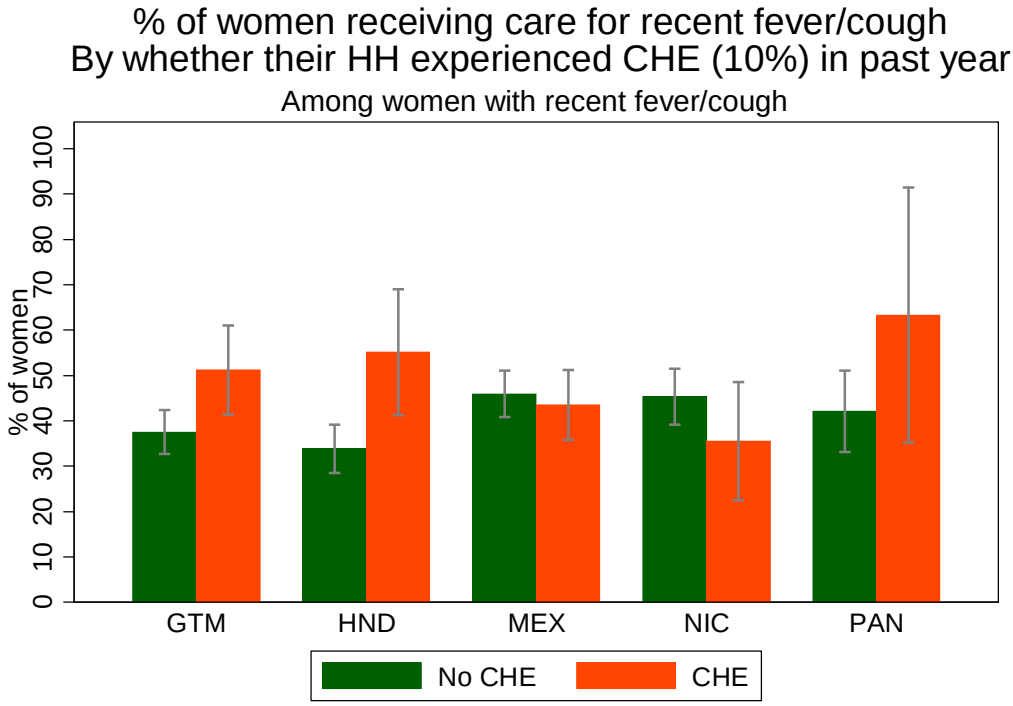
Supplemental Figure S3a



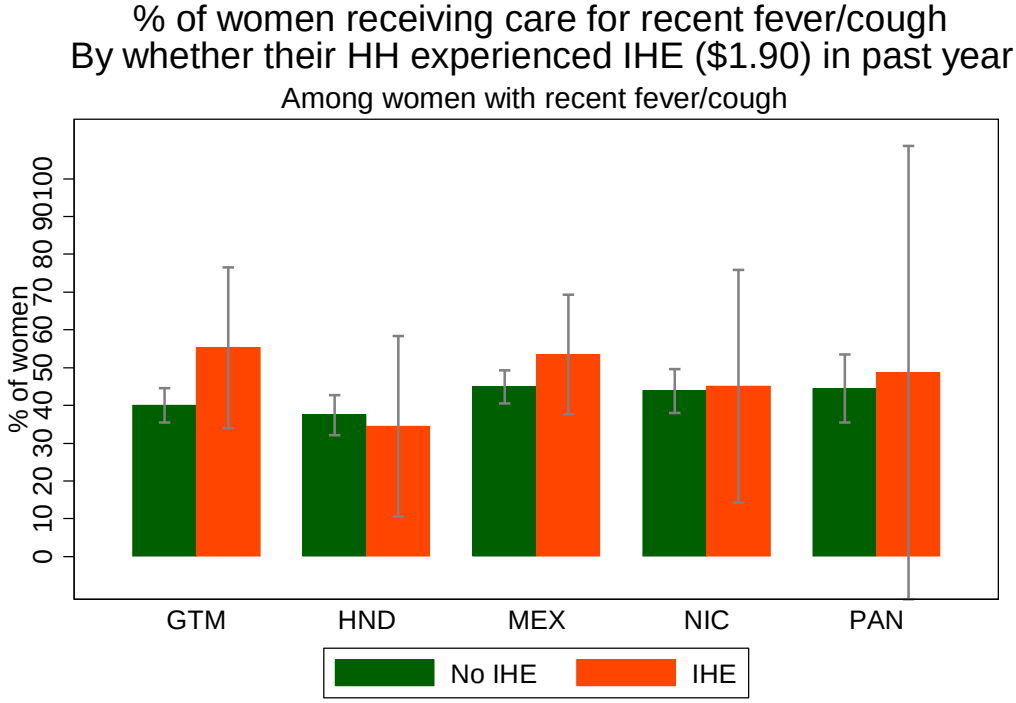
Supplemental Figure S3b



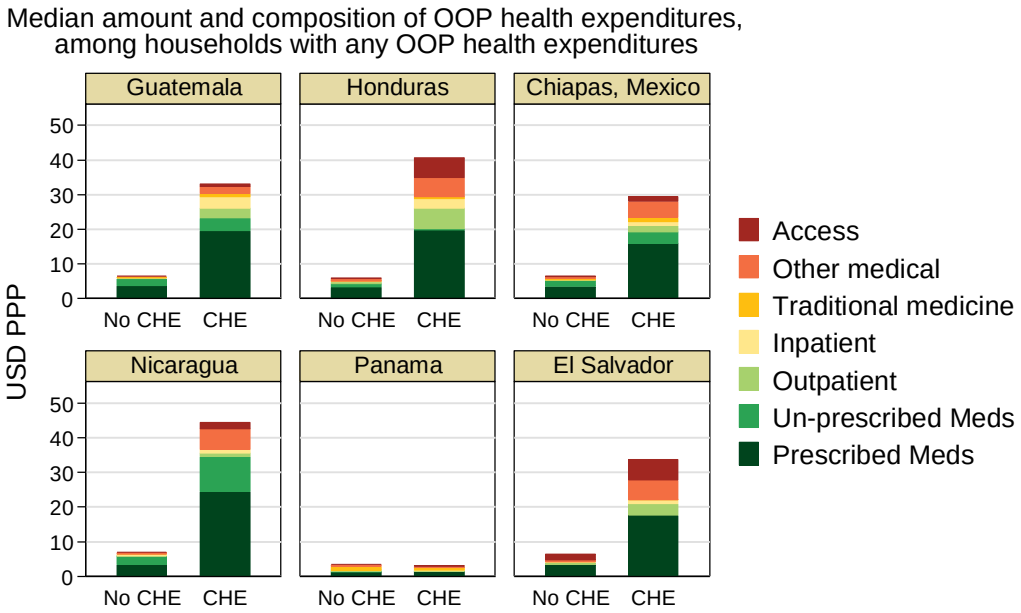
Supplemental Figure S3c



Supplemental Figure S3d

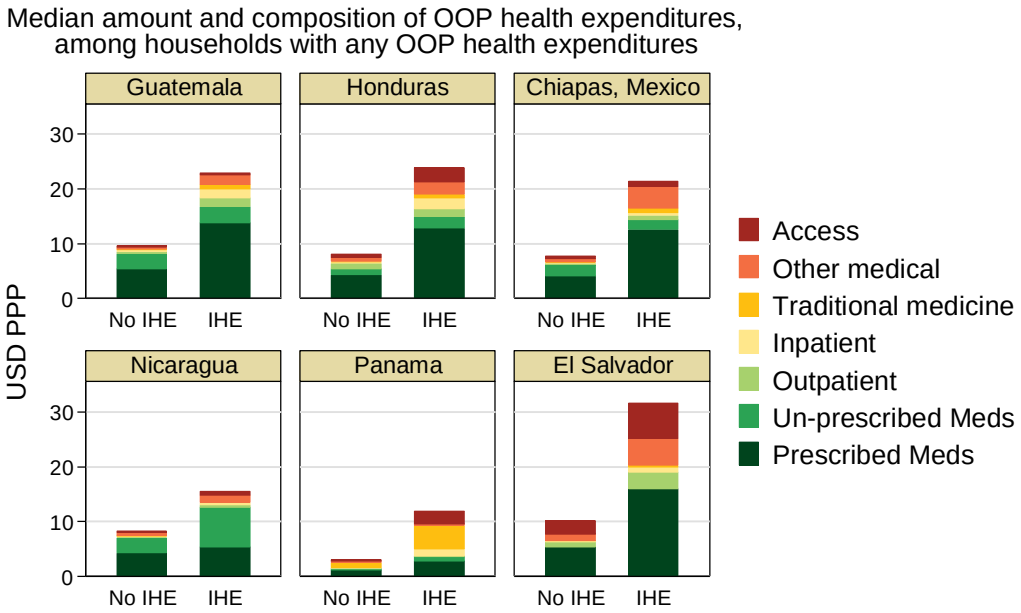


Supplemental Figure S4a



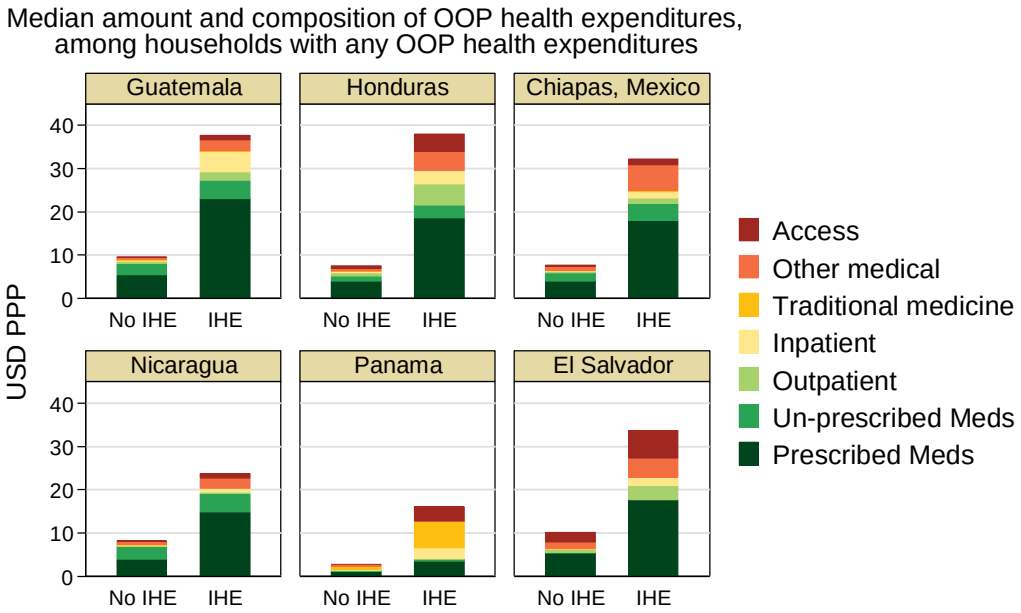
Bars compare households that did or did not experience CHE (25% threshold). Bar height indicates the median OOP health expenditure less insurance reimbursements. Colors within bars indicate the average proportional distribution of different types of medical expenditures.

Supplemental Figure S4b



Bars compare households that did or did not experience IHE ($1.90 threshold). Bar height indicates the median OOP health expenditure less insurance reimbursements. Colors within bars indicate the average proportional distribution of different types of medical expenditures.

Supplemental Figure S4c



Bars compare households that did or did not experience IHE ($3.10 threshold). Bar height indicates the median OOP health expenditure less insurance reimbursements. Colors within bars indicate the average proportional distribution of different types of medical expenditures.

Supplemental Table S1: Household factors associated with spending over 10% of total expenditure on health in study population, additionally controlling for different types of health expenditures

	Guatemala		Honduras		Chiapas, Mexico		Nicaragua		Panama		El Salvador	
Urban	1.48	[0.74,2.97]	0.76	[0.35,1.63]	0.83	[0.53,1.31]	0.38	[0.14,1.02]	NA		0.86	[0.55,1.33]
HoH language[a]												
Spanish only	Ref		NA		Ref		Ref				NA	
Spanish and indigenous	1.33	[0.63,2.78]	NA		1.03	[0.72,1.46]	0.99	[0.25,3.84]	Ref		NA	
Indigenous only	1.08	[0.31,3.76]	NA		0.93	[0.45,1.94]			4.12*	[1.39,12.23]	NA	
Household size	0.93	[0.86,1.00]	0.93	[0.84,1.04]	0.93*	[0.87,0.99]	0.95	[0.85,1.06]	1.18*	[1.02,1.37]	0.92*	[0.85,0.99]
Female HoH	0.83	[0.52,1.32]	1.18	[0.70,1.97]	0.93	[0.52,1.66]	1.62	[0.86,3.07]	0.51	[0.15,1.73]	1.03	[0.73,1.46]
Any infant <1	0.96	[0.68,1.35]	0.92	[0.59,1.43]	0.95	[0.70,1.28]	1.11	[0.58,2.12]	0.58	[0.28,1.22]	1.05	[0.71,1.57]
Any elderly 60+	1.01	[0.69,1.49]	1.07	[0.64,1.80]	1.64*	[1.05,2.57]	1.46	[0.87,2.44]	0.86	[0.29,2.52]	1.28	[0.84,1.94]
Travel time to usual health facility												
<15 min	Ref		Ref		Ref		Ref		Ref		Ref	
15-29 min	1.07	[0.54,2.15]	1.06	[0.66,1.70]	1.04	[0.69,1.58]	1.68	[0.74,3.84]	1.37	[0.25,7.56]	1.19	[0.75,1.88]
30-59 min	1.04	[0.57,1.89]	1.32	[0.71,2.45]	1.23	[0.81,1.88]	1.03	[0.48,2.21]	0.46	[0.05,4.48]	1.04	[0.65,1.65]
60+ min	0.72	[0.37,1.40]	1.16	[0.65,2.04]	1.65*	[1.01,2.67]	1.35	[0.63,2.87]	1.41	[0.54,3.67]	0.9	[0.52,1.57]
HH member hospitalized in past year	1.99	[0.73,5.41]	0.94	[0.51,1.73]	0.68	[0.40,1.18]	1.04	[0.60,1.80]	0.93	[0.20,4.21]	NC	
Has health insurance	0.74	[0.47,1.16]	NA		1.03	[0.67,1.58]	0.65	[0.32,1.33]	0.53	[0.07,3.94]	NA	
Enrolled in opportunidades	NA		NA		1.3	[0.94,1.82]	NA		NA		NA	
Maternal education (highest in HH)												
None	Ref		Ref		Ref		Ref		Ref		Ref	
Primary	1.69	[0.97,2.95]	1.53	[0.58,4.03]	1.21	[0.74,1.97]	0.76	[0.19,3.16]	2.2	[0.58,8.33]	0.8	[0.45,1.43]
Post-Primary	2.47**	[1.37,4.47]	1.38	[0.53,3.58]	0.99	[0.59,1.69]	1.32	[0.33,5.38]	0.64	[0.17,2.44]	0.87	[0.50,1.50]
Wealth quintile												
1 (Poorest)	Ref		Ref		Ref		Ref		Ref		Ref	
2	0.54	[0.27,1.05]	0.54	[0.25,1.17]	1.33	[0.82,2.15]	1.2	[0.48,2.99]	1.89	[0.42,8.45]	0.53*	[0.31,0.91]
3	0.55	[0.27,1.13]	0.99	[0.49,2.03]	1.04	[0.67,1.61]	1.36	[0.46,3.99]	0.8	[0.16,4.08]	0.82	[0.49,1.37]
4	0.68	[0.33,1.42]	0.87	[0.43,1.75]	0.75	[0.47,1.20]	0.92	[0.37,2.28]	1.11	[0.23,5.39]	1	[0.59,1.67]
5 (Wealthiest)	0.64	[0.29,1.40]	1.15	[0.55,2.40]	0.61	[0.34,1.12]	0.64	[0.19,2.18]	0.92	[0.11,7.54]	0.82	[0.47,1.43]
Any inpatient expenditure	1.04*	[1.01,1.08]	1	[0.97,1.03]	1.04**	[1.01,1.06]	1.04*	[1.01,1.07]	1.05**	[1.01,1.08]	1.04***	[1.03,1.05]
Any outpatient expenditure	1.05**	[1.02,1.08]	1.01*	[1.00,1.03]	1.02***	[1.01,1.03]	1.02*	[1.00,1.03]	1.02*	[1.00,1.04]	1.03***	[1.02,1.04]

Any traditional medicine expenditure	1.01	[0.99,1.04]	0.96	[0.88,1.04]	1.02	[1.00,1.05]	1	[1.00,1.00]	1.03***	[1.01,1.04]	1.01	[0.99,1.03]
Any other medical expenditure	1.02*	[1.00,1.03]	1.03***	[1.02,1.05]	1.05***	[1.04,1.06]	1.04***	[1.02,1.06]	1.02	[0.99,1.05]	1.04***	[1.03,1.04]
Any medication expenditure	1.06***	[1.05,1.06]	1.04***	[1.04,1.05]	1.06***	[1.05,1.06]	1.05***	[1.04,1.05]	1.04***	[1.02,1.05]	1.04***	[1.04,1.05]
Any access expenditure	1.01	[0.99,1.03]	1.02**	[1.01,1.04]	1.03***	[1.01,1.05]	1.04	[0.99,1.08]	1.03	[1.00,1.05]	1.02***	[1.01,1.03]
Any lost income expenditure	1.01*	[1.00,1.02]	1.01*	[1.00,1.02]	1.01**	[1.00,1.01]	1	[1.00,1.01]	0.98	[0.96,1.01]	NC	
N	3878		2661		5027		1980		1279		3466	

NA: Not applicable, NC: Not collected. [a]In Panama, no households that speak only Spanish (3% of households) had a CHE; we combined Spanish-only speaking with the Spanish and indigenous-speaking category in the regression analysis and compared with the indigenous-only category.

Supplementary Table S2: Household factors associated with spending over 25% of total expenditure on health in study population

	Guatemala		Honduras		Chiapas, Mexico		Nicaragua		Panama		El Salvador	
Urban	1.50	[0.74,3.02]	1.20	[0.71,2.02]	1.03	[0.72,1.47]	0.50	[0.24,1.04]	NA		1.27	[0.87,1.86]
HoH language[a]												
Spanish only	Ref		NA		Ref		Ref				NA	
Spanish and indigenous	1.91*	[1.16,3.12]	NA		0.77	[0.58,1.03]	0.30	[0.05,1.81]	Ref		NA	
Indigenous only	0.78	[0.25,2.40]	NA		0.82	[0.48,1.40]			4.46*	[1.30,15.30]	NA	
Household size	1.01	[0.94,1.08]	0.93	[0.83,1.03]	1.02	[0.97,1.08]	0.88	[0.77,1.00]	1.19	[0.97,1.45]	0.95	[0.91,1.00]
Female HoH	0.95	[0.55,1.63]	0.84	[0.46,1.51]	0.96	[0.62,1.50]	0.92	[0.45,1.88]	1.54	[0.81,2.90]	0.98	[0.75,1.29]
Any infant <1	0.99	[0.70,1.41]	1.08	[0.62,1.88]	0.96	[0.74,1.25]	0.98	[0.54,1.78]	1.19	[0.57,2.46]	1.34	[0.99,1.81]
Any elderly 60+	1.10	[0.74,1.63]	1.77*	[1.15,2.73]	1.26	[0.83,1.90]	0.86	[0.37,1.99]	0.70	[0.29,1.70]	1.51**	[1.15,1.98]
Travel time to usual health facility												
<15 min	Ref		Ref		Ref		Ref		Ref		Ref	
15-29 min	0.83	[0.51,1.33]	0.70	[0.38,1.29]	0.94	[0.63,1.40]	1.13	[0.56,2.27]	0.75	[0.15,3.79]	1.10	[0.79,1.55]
30-59 min	0.98	[0.56,1.71]	0.74	[0.37,1.48]	0.95	[0.65,1.38]	0.60	[0.28,1.29]	0.89	[0.23,3.47]	1.26	[0.87,1.83]
60+ min	1.22	[0.75,2.00]	1.07	[0.63,1.84]	1.26	[0.78,2.03]	0.68	[0.32,1.44]	1.19	[0.43,3.28]	1.31	[0.79,2.18]
HH member hospitalized in past year	6.94***	[4.41,10.91]	3.00***	[1.82,4.97]	2.93***	[2.09,4.12]	1.86*	[1.03,3.34]	2.00	[0.67,6.00]	NC	[1.00,1.00]
Has health insurance	2.52***	[1.75,3.61]	NA		0.93	[0.63,1.38]	0.85	[0.36,2.00]	0.35	[0.10,1.20]	NA	[1.00,1.00]
Enrolled in opportunidades	NA		NA		1.00	[0.73,1.36]	NA		NA		NA	[1.00,1.00]
Maternal education (highest in HH)												
None	Ref		Ref		Ref		Ref		Ref		Ref	
Primary	1.10	[0.71,1.72]	1.43	[0.50,4.09]	1.08	[0.72,1.62]	0.70	[0.15,3.19]	3.67	[0.55,24.41]	1.20	[0.68,2.10]
Post-Primary	1.61	[0.90,2.90]	1.56	[0.55,4.39]	1.23	[0.77,1.98]	1.14	[0.24,5.37]	1.63	[0.29,9.11]	1.49	[0.85,2.60]
Wealth quintile												
1 (Poorest)	Ref		Ref		Ref		Ref		Ref		Ref	
2	1.30	[0.77,2.19]	1.14	[0.51,2.52]	1.09	[0.68,1.74]	1.35	[0.47,3.86]	12.74*	[1.12,145.47]	0.91	[0.60,1.37]
3	1.03	[0.64,1.64]	1.61	[0.81,3.19]	0.92	[0.58,1.47]	1.40	[0.47,4.14]	4.39	[0.85,22.62]	0.90	[0.60,1.36]
4	1.09	[0.60,1.97]	1.45	[0.65,3.22]	1.10	[0.69,1.75]	1.43	[0.45,4.52]	9.66	[0.82,113.24]	1.20	[0.81,1.78]
5 (Wealthiest)	0.98	[0.51,1.86]	1.95	[0.97,3.91]	1.25	[0.76,2.03]	3.00	[0.81,11.15]	6.98	[0.52,93.20]	1.56*	[1.01,2.40]
N	3878		2661		5027		1980		1279		3466	

NA: Not applicable, NC: Not collected. [a]In Panama, no households that speak only Spanish (3% of households) had a CHE; we combined Spanish-only speaking with the Spanish and indigenous-speaking category in the regression analysis and compared with the indigenous-only category.

Supplementary Table S3: Household factors associated with impoverishing health expenditure at the $1.90 threshold, in study population

	Guatemala		Honduras		Chiapas, Mexico		Nicaragua		Panama		El Salvador	
Urban	1.15	[0.57,2.31]	0.96	[0.41,2.22]	1.26	[0.79,2.01]	0.49	[0.18,1.35]	NA		1.15	[0.71,1.88]
HoH language[a]												
Spanish only	Ref		NA		Ref		Ref				NA	
Spanish and indigenous	1.27	[0.71,2.27]	NA		0.83	[0.54,1.27]	0.27	[0.04,1.94]	Ref	[1.00,1.00]	NA	
Indigenous only	1.00	[1.00,1.00]	NA		0.87	[0.39,1.90]			4.20*	[1.29,13.64]	NA	
Household size	1.05	[0.97,1.13]	1.01	[0.90,1.15]	1.01	[0.92,1.11]	0.93	[0.81,1.07]	0.96	[0.72,1.28]	0.94	[0.86,1.02]
Female HoH	0.64	[0.28,1.44]	1.02	[0.46,2.26]	0.80	[0.44,1.44]	1.23	[0.55,2.73]	2.95	[0.73,11.91]	1.01	[0.71,1.43]
Any infant <1	0.95	[0.62,1.47]	1.18	[0.59,2.37]	0.77	[0.54,1.10]	1.61	[0.83,3.11]	0.67	[0.14,3.10]	1.48*	[1.03,2.13]
Any elderly 60+	1.01	[0.56,1.83]	1.46	[0.65,3.31]	0.75	[0.39,1.44]	1.04	[0.36,2.97]	0.96	[0.22,4.25]	1.07	[0.69,1.67]
Travel time to usual health facility												
<15 min	Ref		Ref		Ref		Ref		Ref		Ref	
15-29 min	0.67	[0.38,1.16]	0.67	[0.22,2.04]	0.80	[0.49,1.30]	1.59	[0.64,3.94]	6.53	[0.60,70.97]	1.11	[0.71,1.73]
30-59 min	0.66	[0.32,1.35]	0.92	[0.36,2.39]	1.01	[0.60,1.71]	2.08	[0.66,6.53]	2.38	[0.11,52.71]	0.93	[0.54,1.58]
60+ min	0.90	[0.49,1.68]	1.51	[0.66,3.47]	1.22	[0.71,2.10]	1.05	[0.40,2.78]	2.97	[0.57,15.41]	1.76	[0.92,3.40]
HH member hospitalized in past year	3.19**	[1.49,6.81]	3.46***	[1.85,6.48]	2.91***	[1.76,4.83]	1.44	[0.58,3.59]	1.37	[0.16,11.36]	NC	[1.00,1.00]
Has health insurance	2.10**	[1.27,3.45]	NA		1.66	[0.83,3.31]	1.35	[0.50,3.65]	0.42	[0.04,3.87]	NA	[1.00,1.00]
Enrolled in opportunidades	NA		NA		1.36	[0.87,2.14]	NA		NA		NA	
Maternal education (highest in HH)												
None	Ref		Ref		Ref		Ref		Ref		Ref	
Primary	1.30	[0.67,2.50]	0.94	[0.29,3.03]	0.79	[0.47,1.32]	0.58	[0.20,1.69]	3.97	[0.36,43.50]	1.04	[0.48,2.27]
Post-Primary	1.45	[0.68,3.12]	1.61	[0.40,6.44]	0.81	[0.46,1.42]	0.74	[0.19,2.91]	0.64	[0.01,29.01]	1.16	[0.51,2.64]
Wealth quintile												
1 (Poorest)	Ref		Ref		Ref		Ref		Ref		Ref	
2	0.74	[0.33,1.65]	0.59	[0.22,1.63]	1.16	[0.63,2.13]	0.86	[0.22,3.46]	2.69	[0.24,30.08]	1.28	[0.68,2.40]
3	0.81	[0.37,1.75]	1.04	[0.44,2.48]	0.94	[0.47,1.90]	1.65	[0.43,6.28]	7.64	[0.81,71.91]	1.21	[0.62,2.35]
4	1.04	[0.44,2.43]	0.69	[0.24,1.98]	2.29**	[1.23,4.25]	2.34	[0.70,7.85]	0.94	[0.07,12.82]	1.85	[0.91,3.76]
5 (Wealthiest)	0.76	[0.31,1.88]	0.67	[0.22,2.03]	1.44	[0.63,3.28]	1.73	[0.29,10.27]	5.20	[0.76,35.53]	1.82	[0.94,3.55]
N	3618		2661		5027		1980		1279		3466	

NA: Not applicable, NC: Not collected. [a]In Panama, no households that speak only Spanish (3% of households) had a CHE; we combined Spanish-only speaking with the Spanish and indigenous-speaking category in the regression analysis and compared with the indigenous-only category.

Supplementary Table S4: Household factors associated with impoverishing health expenditure at the $3.10 threshold, in study population

	Guatemala		Honduras		Chiapas, Mexico		Nicaragua		Panama		El Salvador	
Urban	1.01	[0.39,2.62]	1.63	[0.73,3.65]	0.97	[0.56,1.67]	0.43	[0.14,1.28]	NA		1.26	[0.74,2.14]
HoH language[a]												
Spanish only	Ref		NA		Ref		Ref				NA	
Spanish and indigenous	2.81**	[1.34,5.87]	NA		0.61	[0.36,1.02]	1.04	[0.27,4.09]	Ref		NA	
Indigenous only	1.22	[0.15,9.55]	NA		0.27	[0.06,1.25]			4.69**	[1.85,11.93]	NA	
Household size	0.93	[0.86,1.01]	0.88	[0.75,1.02]	0.93	[0.82,1.06]	1.00	[0.87,1.14]	0.91	[0.72,1.15]	0.85***	[0.78,0.93]
Female HoH	0.77	[0.34,1.75]	0.98	[0.50,1.94]	1.21	[0.65,2.23]	0.63	[0.24,1.67]	2.16	[0.62,7.51]	1.00	[0.69,1.45]
Any infant <1	0.74	[0.41,1.31]	1.31	[0.69,2.48]	1.01	[0.68,1.51]	1.77	[0.74,4.24]	0.17*	[0.04,0.68]	1.28	[0.79,2.08]
Any elderly 60+	0.51	[0.22,1.20]	1.40	[0.68,2.90]	1.27	[0.71,2.28]	0.58	[0.12,2.83]	1.00	[0.35,2.88]	1.12	[0.70,1.78]
Travel time to usual health facility												
<15 min	Ref		Ref		Ref		Ref		Ref		Ref	
15-29 min	1.01	[0.54,1.91]	0.58	[0.23,1.45]	1.08	[0.64,1.81]	1.58	[0.48,5.20]	2.48	[0.54,11.43]	1.84*	[1.06,3.19]
30-59 min	0.80	[0.33,1.92]	0.48	[0.16,1.45]	1.01	[0.59,1.73]	1.30	[0.38,4.51]	0.28	[0.07,1.20]	1.30	[0.65,2.58]
60+ min	1.10	[0.47,2.61]	0.68	[0.28,1.65]	1.06	[0.57,1.96]	1.18	[0.30,4.69]	0.87	[0.20,3.89]	1.59	[0.75,3.36]
HH member hospitalized in past year	7.02***	[3.94,12.53]	2.22*	[1.05,4.68]	2.45***	[1.48,4.05]	2.58*	[1.20,5.55]	5.41**	[1.50,19.54]	NC	
Has health insurance	1.14	[0.63,2.07]	NA		1.52	[0.65,3.59]	0.58	[0.17,2.02]	0.25	[0.03,2.45]	NA	
Enrolled in opportunidades	NA		NA		1.40	[0.84,2.34]	NA		NA		NA	
Maternal education (highest in HH)												
None	Ref		Ref		Ref		Ref		Ref		Ref	
Primary	0.58	[0.31,1.08]	0.97	[0.49,1.91]	1.08	[0.52,2.25]	5.35	[0.63,45.39]	0.18*	[0.03,0.97]	0.89	[0.37,2.14]
Post-Primary	1.42	[0.77,2.62]	1.00	[1.00,1.00]	1.30	[0.61,2.76]	3.98	[0.45,35.38]	0.16*	[0.03,0.78]	1.45	[0.56,3.77]
Wealth quintile												
1 (Poorest)	Ref		Ref		Ref		Ref		Ref		Ref	
2	1.58	[0.69,3.58]	1.21	[0.37,3.97]	1.23	[0.42,3.60]	1.60	[0.35,7.23]	1.65	[0.42,6.56]	0.81	[0.39,1.67]
3	0.84	[0.28,2.51]	1.53	[0.54,4.33]	1.95	[0.72,5.23]	2.18	[0.52,9.17]	1.06	[0.21,5.43]	0.93	[0.49,1.80]
4	1.61	[0.60,4.33]	1.56	[0.53,4.54]	3.33*	[1.26,8.78]	2.23	[0.67,7.42]	0.62	[0.10,3.69]	1.30	[0.64,2.65]
5 (Wealthiest)	1.76	[0.71,4.37]	2.00	[0.66,5.99]	3.88**	[1.46,10.33]	4.96*	[1.35,18.26]	1.00	[1.00,1.00]	1.81	[0.99,3.31]
N	3878		2502		5027		1980		1013		3466	

NA: Not applicable, NC: Not collected. [a]In Panama, no households that speak only Spanish (3% of households) had a CHE; we combined Spanish-only speaking with the Spanish and indigenous-speaking category in the regression analysis and compared with the indigenous-only category.

Supplementary Table S5: National estimates of CHE and IHE in the study countries, from national-level surveys

Country	Year	CHE 10%	CHE 25%	IHE $1.90	IHE $3.10
Guatemala	2014	1.36%	0.04%	0.29%	0.22%
	2011	9.36%	2.25%	0.12%	0.80%
	2006	3.48%	1.15%	0.16%	0.42%
	2000	14.48%	2.33%	0.40%	1.03%
Honduras	1998	3.45%	0.43%	NC	NC
Mexico	2012	7.13%	1.91%	0.28%	0.69%
	2010	3.59%	0.75%	0.23%	0.48%
	2008	6.03%	1.51%	0.23%	1.00%
	2004	9.89%	2.74%	0.45%	1.31%
	2002	9.39%	2.20%	0.64%	1.41%
	2000	9.80%	2.44%	0.78%	1.46%
	1998	8.98%	2.05%	0.85%	1.13%
	1996	11.39%	2.77%	0.88%	1.89%
	1994	10.01%	2.59%	0.95%	1.42%
Nicaragua	2014	27.74%	8.89%	3.08%	5.20%
	2009	16.18%	2.47%	1.35%	2.45%
	2001	15.95%	3.23%	1.84%	2.98%
	1998	12.35%	2.09%	1.87%	2.47%
	1993	9.27%	2.20%	1.07%	1.61%
Panama	2008	1.41%	0.22%	0.04%	0.04%
	1997	9.25%	1.84%	0.53%	0.84%

Aim 2: Impact of the results-based aid Salud Mesoamerica Initiative, phases one and two, on financial protection indicators in poor regions of four middle-income countries, and analysis of potential mediating factors

ABSTRACT

Background: Financial protection from catastrophic health expenditures is a core health systems goal and a key element of Universal Health Coverage (UHC). Households in the poorest areas of Nicaragua, Honduras, Guatemala and Chiapas, Mexico are vulnerable to out of pocket health expenditures, with prescription drugs as the largest driver. We tested whether the first two phases of the Salud Mesoamerica Initiative results based aid program indirectly influenced financial protection measures through its efforts to improve the availability and quality of services in public facilities. **Methods:** Municipalities were randomized (Honduras) or matched based on socioeconomic characteristics (Guatemala, Chiapas, Nicaragua) to receive SMI interventions and incentives or serve as comparison areas. At baseline and follow-up, we selected stratified random samples of households and health facilities in intervention and comparison areas to participate in surveys. Primary outcomes included proportion of total household consumption on health overall, and on prescription drugs, and whether the household forwent health care due to cost in the past year. Secondary outcomes included whether the household made any expenditure on health overall, on prescription drugs, and on non-prescription drugs; and the amount of expenditure on each. After matching households on potential confounding characteristics, we used a difference-in-differences framework to measure the impact of SMI on each outcome. We tested whether SMI's impact on the outcomes was mediated by utilization of the public vs. private sector, recent illness in women/children, hospitalization of household members, and availability of drugs at recently visited health facilities. **Findings:** After matching, we analyzed data from 6,305, 5,019, 7,114 and 3,837 households in Guatemala, Honduras, Chiapas and Nicaragua, respectively. In Honduras, the adjusted difference-in-differences analysis found that SMI was statistically significantly associated with reductions in the proportion of total household consumption on prescription drugs (Coefficient (Coef): -0.02, 95% Confidence Interval (CI): -0.03, -0.00), households forgoing care due to cost (Coef: -0.10, 95% CI: -0.20, -0.01), households making any health expenditure (Coef: -0.10, 95% CI: -0.20, -0.01), and households making any prescription drug expenditure (Coef: -0.11, 95% CI: -0.19, -0.02). SMI was not associated with a significant change in the financial protection outcomes in Guatemala, Chiapas or Nicaragua, with the exception of select improvements among Guatemalan and Nicaraguan households that had visited a health facility in the past month. SMI's significant impacts in Honduras were mediated in part by fewer households where a woman or child was self-reported as sick in the past two weeks in intervention areas. SMI significantly improved drug availability in several countries, but this did not mediate SMI's impacts on financial protection. **Interpretation:** There are early indications that SMI's systems strengthening approach can feasibly improve financial outcomes through the pathways of increasing utilization of the private sector and reducing illness, particularly in Honduras. However, SMI's effects are inconsistent depending on setting. Additional research is needed to understand the contextual factors determining why SMI improved financial protection in some countries but not others.

INTRODUCTION

Financial protection from catastrophic or impoverishing health expenditures is a core health systems goal and a key element of Universal Health Coverage (UHC). In the Latin America and Caribbean Region, 13.4% of households spend over 10% of total consumption on health, more than any other region in the world [1]. Households in the poorest quintile are the most vulnerable and can experience several times higher incidence of catastrophic health expenditure than those in the wealthiest quintile [2]. Despite many regional and national efforts to promote UHC and financial protection, pockets of vulnerable and underserved populations remain [3, 4].

To address health disparities in their most marginalized areas, Guatemala, Honduras, Nicaragua, the State of Chiapas, Mexico and three other countries began the Salud Mesoamerica Initiative (SMI) results-based aid (RBA) program in 2011 [5]. Under SMI, each Ministry of Health (MOH) committed to improving the supply, quality and coverage of maternal and child health services in regions with a high concentration of households in the poorest quintile. The MOHs then negotiated targets for process and outcome indicators with the Interamerican Development Bank (IDB) and donors; meeting these targets, as assessed by an independent evaluator, was tied to subsequent payments.

Prior to the implementation of SMI, 8.9%, 9.7%, 9.9% and 13.8% of households in the poor areas targeted by SMI in Nicaragua, Honduras, Guatemala and Chiapas, respectively, spent over 10% of total consumption on health [Source: Aim 1]. Mirroring global patterns, the largest driver of health expenditures in these poor populations was medications, and specifically prescription drugs (as opposed to those purchased without a prescription), which accounted for between 54% (Nicaragua) and 67% (Guatemala) of health expenditures in households that spent over 10% of total consumption on health [6, 7]. The extent of expenditures on health and prescription drugs is discordant with policies promising free or low-cost services and essential medicines at public facilities in Guatemala, Honduras, Nicaragua and Mexico [8–10]. Furthermore, more than one in five households in SMI areas reported forgoing care due to cost in the past year in all four of these countries [Source: Aim 1].

While financial protection outcomes are not tied to any SMI payments, they are tracked as monitoring indicators and could be influenced by SMI through a number of potential mechanisms. First, it is important to monitor if increased coverage and utilization of services has the unintended consequence of driving up costs as more people receive services. Conversely, there are several pathways by which SMI could improve financial protection, for instance by reducing the need for expensive curative care through increased coverage of preventive services. Additionally, previous studies show that households in SMI areas take quality into account when selecting a facility [11]. SMI has worked to improve the quality of public facilities; if successful, this could lead households to choose low-cost public options over paying private providers. Finally, qualitative research in SMI areas of Chiapas, Mexico suggests that prescription drug expenditures occur when public facilities stock-out of free/low-cost drugs, forcing household to purchase elsewhere [5, 10, 12]. These SMI countries have payment indicators tied to reducing drug stock-outs, offering another possible mechanism for improving financial protection.

This study aims to evaluate the impact of SMI on financial protection indicators in Guatemala, Honduras, Nicaragua and Chiapas, Mexico, and test several mechanisms hypothesized to mediate this impact.

METHODS

Initiative and study design

The SMI methodology has been previously published, including details on the design, survey instruments, sampling and data collection [5]. SMI targets the poorest regions of Guatemala, Honduras, Nicaragua and the state of Chiapas, Mexico. In 2011, the MOH of each country worked with the IDB to identify their departments (or municipalities in the case of Chiapas) with the highest concentration of households in the poorest quintile. Municipalities were selected from within these departments to serve as either SMI intervention or comparison areas. In Honduras, selection and assignment of municipalities was random, while in Guatemala, Chiapas and Nicaragua, assignment was done by matching intervention and comparison municipalities with similar socioeconomic profiles.

For municipalities selected for SMI interventions, the MOH and IDB negotiated specific targets for process and outcome measures related to maternal and child health. The MOH was then responsible (with support from IDB and technical assistance partners) for designing and implementing interventions to achieve these outcomes in SMI intervention municipalities. The comparison municipalities were not part of target setting and did not receive any financial incentives based on performance. Some SMI interventions were targeted only to SMI areas, such as support for monitoring and continuous quality improvement. However, other interventions initiated under SMI could not feasibly be targeted to only the comparison areas, including changes to country norms.

SMI was designed with three phases. The first phase, taking place 2012-2014, focused on increasing the availability of essential drugs and equipment. The second phase, taking place 2015-2017, centered on improving service provision and quality. The third phase, upcoming, focuses on increasing demand for services. This study uses data collected prior to the first phase (baseline), and following the second phase (follow up).

At baseline and following each phase, an independent evaluator (the authors of this study) conducted household and health facility surveys to measure progress against the pre-negotiated indicator targets and determine if the incentive-based payments would be dispersed. In Guatemala, Honduras, Nicaragua and Chiapas, these surveys were conducted in both intervention and comparison areas, to establish a counterfactual. However, the payments were tied only to meeting targets in the intervention areas, regardless of what occurred in the comparison areas.

Sampling method

Stratified random samples of households and health facilities were selected at baseline and follow-up (following the end of SMI's 2nd phase). Baseline and follow-up samples were selected separately and sample sizes were powered to detect changes in the SMI payment indicators [5].

Household sample: In each country, the household sampling frame included all census segments (our primary sampling unit) located in SMI and comparison areas, from the most recent national census. We randomly sampled segments within the SMI and comparison strata, using probability proportional to the number of households. In each selected segment, we conducted a new household census to account for

population movement since the last government census. We then identified eligible households based on our census and randomly selected 30 eligible households in each segment to be part of the survey. Household samples were selected independently at baseline and follow-up to avoid introducing observer bias by re-visiting the same households or communities; the same procedure was used to select the baseline and follow-up samples. Inclusion criteria to participate in the household survey were households with a woman aged 15-49, or a child under five.

Health facility sample: The health facility frame for each country included all public facilities in SMI intervention and comparison municipalities, stratified by their Essential Obstetric and Neonatal Care rating (EONC). At baseline, facilities were randomly selected from each of the EONC strata for intervention and comparison areas. At follow-up, we further stratified between health facilities that had and had not been surveyed at baseline, and selected approximately half of the sample from each stratum. Within each stratum, we first selected from among facilities that were in a municipality included in the household sample and only selected facilities from other municipalities if it was needed to meet the sample quota. This was done to allow for linked facility-household analyses.

Survey instruments

Household and health facility survey instruments were designed in collaboration with content matter experts from the IDB and each MOH. They were consistent across countries, with some exceptions (e.g. national insurance programs). The instruments underwent piloting and revisions before data collection.

This household survey contained a household census, household module, woman module, and child module. The census collected information on the gender, age and relationship of each resident. The household module included questions on total expenditure, expenditure on non-health items (food, drinks, tobacco, utilities, clothing, transport, communication, school supplies, household maintenance, hobbies, and other), and expenditure on health (total, prescribed drugs, un-prescribed drugs, outpatient fees, inpatient fees, traditional healers, dentists, other medical supplies, lab tests, ambulance fees, other health-related expenses such as transport, and the value of in-kind payments). Households reported if any member had forgone care due to cost in the past year. The woman module, each woman reported their education level, insurance status, participation in social benefit programs, recent illness, type of facility usually visited, if they visited a health facility in the past month and name of that facility. Finally, the child module contained questions about recent illness.

In the health facility survey, the enumerator recorded whether a list of tracer medications for maternal and child health were observed to be available on the day of the survey (among other questions). Our primary measures of drug availability used information on a set of tracer drugs were specific to the facility's EONC level, and common across countries[1]. In addition, each country's survey contained additional questions on country-specific drugs tied to payment indicators (Supplementary Table 1). For

[1] By EONC level, the list of trader drugs included six items for ambulatory facilities without a doctor [ferrous sulfate or multivitamin; oral rehydration solution; albendazole or mebendazole; folic acid; zinc sulfate or gluconate; iron], seven for ambulatory facilities with a doctor [same as ambulatory without a doctor, plus an antibiotic], 12 for basic facilities [same as ambulatory with a doctor, plus dexamethasone, magnesium sulfate, oxytocin, hydralazine and gentamycin], and 13 for complete facilities [same as basic facilities, plus diazepam].

these country-specific indicator drugs, information was collected on day-of-availability and whether any stock-out had occurred in the past three months (recorded in documents) for each drug.

Data collection

Baseline surveys were collected between July 2012 and August 2013[2]. Follow-up surveys were collected between May 2017 and October 2018[3]. All data were collected by trained personnel from local data collection agencies using computer-assisted personal interview (CAPI) with the DatStat Illume software. Data were uploaded daily to a central server, allowing continuous data quality monitoring and feedback.

The study received institutional review board (IRB) approval from the University of Washington, partnering data collection agencies, and the MOH of each country. Informed consent was obtained from all women and health facility directors prior to data collection.

Impact of SMI on financial protection outcomes

Standardization of expenditure data

Different expenditure variables were asked for different durations of time (e.g. food expenditure in the past week; prescription drug expenditures over the past month), so we standardize all expenditures to a month-long period. We visually inspected the data to remove extreme outliers, and checked that the sub-components of each expenditure category summed to the reported totals. If the sub-components and reported total could not be reconciled, we opted to use the sum of the sub-components [13]. We standardized all expenditures to 2018 USD purchasing power parity (PPP) using World Bank PPP conversion factors [14] and US government inflation rates [15].

Outcome measures

Primary outcomes included the proportion of total household consumption spent on health (less insurance reimbursements); the proportion of total household consumption spent on prescription drugs (the largest driver of health expenditures); and whether the household forwent care due to cost in the past year.

Secondary outcomes included whether the household reported any health expenditure; the amount of OOP health expenditure (2018 USD PPP, less insurance reimbursements); whether the household made any prescription drug expenditure; the amount of OOP prescription drug expenditure (2018 USD PPP); whether the household made any non-prescription drug expenditure; the amount of OOP non-prescription drug expenditure (2018 USD PPP).

Statistical analysis of impact

We analyzed the impact of SMI on each of the financial protection outcomes using matching combined with a difference-in-differences framework. All analyses accounted for sample weights and clustering of households within segments.

2 Chiapas, Mexico: July 2012-May 2013; Honduras: Jan-June 2013; Nicaragua: Mar-Aug 2013; Guatemala: Apr-Aug 2013
3 Honduras: May-Nov 2017; Nicaragua June-Nov 2017; Chiapas, Mexico: Feb-July 2018; Guatemala: May-Oct 2018

SMI was not designed to collect repeat measurements on individual households. Given that assignment of municipalities to intervention and control was not random in all countries except Honduras, we aimed to improve comparability across time and between intervention and comparison groups by using coarsened exact matching to create a sample of households with similar characteristics. To implement this, we defined strata based on factors previously found to be significantly associated with catastrophic health expenditure in these communities (Aim 1), including household size, ethnicity, insurance status, highest maternal education level in the household and total household consumption amount. Despite being associated with catastrophic expenditures, we did not match on whether someone from the household had been hospitalized in the past month, because we considered this a potential mediating factor influenced by SMI. We excluded households missing any of the matching variables. We then trimmed the dataset to include only households belonging to strata with representation from all four samples (baseline intervention, baseline comparison, follow up intervention, follow up comparison), and recalculated the weights per the method described by Iacus et al [16].

We conducted descriptive analyses on this matched dataset to examine how the outcomes had changed from baseline to follow-up in intervention and comparison areas.

While matching and trimming provide common support across the samples, residual and unobserved confounding can remain. As an example, baseline and follow-up surveys were sometimes collected in different seasons (intervention and comparison areas were surveyed simultaneously). To mitigate this concern and relax the assumptions needed to assert causality, we analyzed the impact of SMI, among the matched households, in a linear difference-in-differences framework:

$$\textit{Financial protection outcome} = \beta_0 + \beta_1 \textit{Follow-up} + \beta_2 \textit{SMI} + \beta_3 \textit{Followup*SMI} + \beta_4 X + \varepsilon$$

Where β_3 is the treatment effect and X is a vector of household covariates (the same variables used for matching, but with household size and total expenditure as continuous rather than coarsened categorical variables). The analysis was conducted separately for each country.

As sensitivity analyses, we restricted to households that had visited a health facility in the month prior to the survey and estimated the crude effect-size without matching or controlling for covariates.

Mediation analysis

For the outcomes and countries where we found SMI had a significant impact, we conducted a mediation analysis to test hypothesized causal pathways between SMI and the outcomes.

Description of hypothesized mediating factors

Shift in care seeking from private to public sector: We hypothesized that by improving the availability and quality of services at public facilities, SMI could have shifted utilization from the private to the public sector, which is typically less expensive. The variable used to examine this potential mediator was a binary indicator of whether the household's self-reported usual health facility was public or private.

Improved health status: We hypothesized that by improving the quality and availability of care (especially preventive care) SMI could have improved health (real or self-perceived) in the communities,

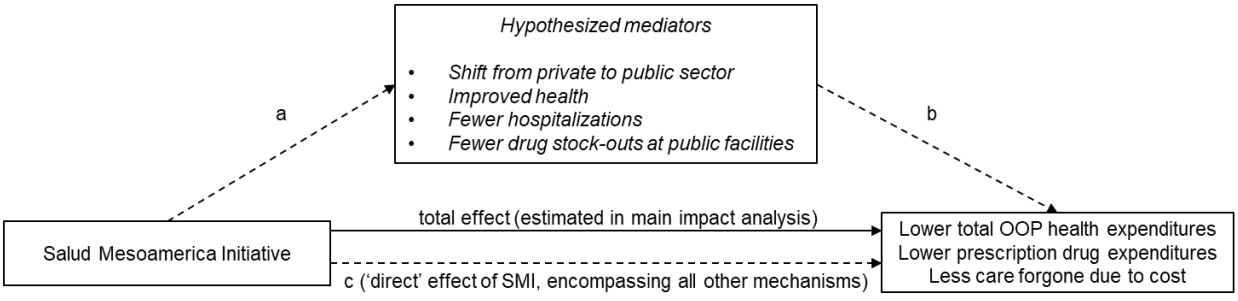
therefore reducing the need to purchase medical care and drugs. The variable used to examine this potential mediator was a binary indicator of whether a woman or child in the household had been sick in the past 2 weeks

Fewer hospitalizations: We hypothesized that by improving the quality and availability of care (especially preventive care), SMI could have reduced hospitalization rates. As hospitalization was associated with greater OOP expenditure at baseline, this could improve financial protection outcomes. The variable used to examine this potential mediator was a binary indicator of whether a member of the household had been hospitalized in the past year.

Fewer drug stock-outs at public facilities: We hypothesized that by reducing drug stock-outs at public facilities, SMI would have better access to free/low-cost public drugs, rather than purchasing from private pharmacies. This could reduce prescription drug expenditures, the largest type of OOP health expenditure in the baseline analysis. This particular mediation analysis was conducted among a subset of households that reported visiting (in the past month) one of the health facilities that was included in our sample for the health facility survey. This allowed us to link the household's financial protection data with information on drug availability at the facility most recently visited by a household member. We re-did the matching and re-calculated the weights on this sub-set of households. Our main measure of drug availability was the percent of standard tracer medicines[4] observed to be in stock at the facility on the day of the survey. We also conducted a sensitivity analysis looking at the availability of country-specific drugs associated with payment indicators.

Statistical analysis of mediating factors

Figure 1: Hypothesized causal mechanism for the impact of the Salud Mesoamerica Initiative on out of pocket prescription drug expenditures, mediated by reduction in drug stock outs



We tested each link of the hypothesized causal pathway and conducted a mediation analysis (Figure 1). All analyses accounted for sampling weights and clustering of households within segments.

[4]

First, we first examined the impact of SMI on each hypothesized mediator ('a' in Figure 1). Following descriptive analysis, we used a difference-in-differences framework to analyze the impact of SMI on the hypothesized mediator:

$$Mediator = \beta_0 + \beta_1 Follow\text{-}up + \beta_2 SMI + \beta_3 Followup*SMI + \beta_4 X + \varepsilon$$

For most mediators, the unit of analysis was the household, and *X* was the same set of covariates controlled for the impact analysis. For drug stock-outs, the unit of analysis was the health facility, and *X* was a covariate controlling for the facility's EONC level.

Second, we examined the cross-sectional associations between the hypothesized mediators and the financial protection outcomes (`b' in Figure 1), using the following regression model:

$$Financial\ protection\ outcome = \beta_0 + \beta_1 Mediator + \beta_3 X + \varepsilon$$

The unit of analysis was the household for all combinations of outcomes and mediators. The analysis was pooled across baseline, follow-up, intervention and comparison, by country. *X* was a vector of the same household covariates; *X* also included EONC level when analyzing the relationship with drug availability.

Finally, we conducted a product of coefficients mediation analysis with the following specification [17]:

$$Mediator = \beta_0 + \beta_1 Follow\text{-}up + \beta_2 SMI + \beta_3 Followup*SMI + \beta_4 X + \varepsilon$$

$$Financial\ protection\ outcome = \beta_{5+} \beta_6 Follow\text{-}up + \beta_7 SMI + \beta_8 Followup*SMI + \beta_9 Mediator + \beta_{10} X + \varepsilon$$

$$B_{indirect} = (\beta_3)*(\beta_9)$$

The average causal mediation effect ($B_{indirect}$) was estimated as the product of $\beta_{3\,(}$the coefficient for the impact of the intervention on the mediator) and β_9 (the coefficient for the impact of the mediator on the outcome, independent of the intervention). The analysis was implemented using the *mediate* package in R [18].

RESULTS

Sample description

Including both baseline and follow up surveys, we collected data from 6,577, 5,097, 7,670 and 4,521 households in Guatemala, Honduras, Chiapas and Nicaragua, respectively (Supplementary Table 2). Of these, 6,305, 5,019, 7,114 and 3,837 remain after matching. A breakdown by intervention vs. comparison and baseline vs. follow-up is shown in Table 1, with characteristics shown in Supplementary Table 3.

We collected data from 129, 120, 128 and 111 ambulatory-level facilities and 56, 49, 52 and 43 basic/complete-level facilities in Guatemala, Honduras, Chiapas and Nicaragua, respectively (Table 1). 59, 57, 65 and 63 of the ambulatory facilities and 49, 43, 49 and 19 of the basic/complete facilities were surveyed at both baseline and follow-up, in Guatemala, Honduras, Chiapas and Nicaragua respectively.

553 (Guatemala), 946 (Honduras), 415 (Chiapas), and 338 (Nicaragua) matched households reported visiting a health facility in the past month that was surveyed as part of our health facility sample. These households were linked to surveys from a total of 71 (Guatemala), 75 (Honduras), 44 (Chiapas) and 44 (Nicaragua) health facilities.

Impact of SMI on financial protection outcomes

Impact of SMI on total OOP health expenditure

At baseline, the mean percent of total consumption spent on health ranged from 2.9% (Nicaragua) to 5.4% (Chiapas) in SMI intervention areas, and was similar in comparison areas of each country (Figure 2A). There were minimal changes in this indicator between baseline and follow-up in SMI intervention areas (maximum change: +0.6 percentage points, in Guatemala), but in comparison areas the average percent of total consumption spent on health increased in Honduras, Chiapas and Nicaragua and decreased in Guatemala (maximum change: +2.8%-points, in Honduras). However, in the difference-in-differences analysis, SMI did not have a statistically significant impact on proportion of total consumption spent on health in any country (Guatemala Coefficient (Coef): -0.00, 95% Confidence Interval (CI): -0.02, 0.02; Honduras Coef: -0.02, 95% CI: -0.05, 0.01; Chiapas Coef: -0.01, 95% CI: -0.04, 0.02; Nicaragua Coef -0.03, 95% CI: -0.08, 0.01) (Table 1). However, we did note that in a sensitivity analyses that used a difference-in-differences framework but did not trim un-matched households or adjust for covariates or clustering of households within sampling units, SMI was associated with a significantly lower proportion spent on health in Honduras only (Coef: -0.03, 95% CI: -0.05, -0.01).

Among secondary outcomes, the percent of households making any OOP health expenditure decreased in SMI areas of Guatemala, Honduras and Nicaragua, and increased in Chiapas, while increasing in comparison areas of Honduras, Chiapas and Nicaragua and decreasing in Guatemala. The largest changes were in Honduras, where the proportion of households making any health expenditure decreased from 18.1% to 12.5% in SMI intervention areas, while increasing 15.0% to 27.0% in comparison areas (Supplementary Figure 1A). In the difference-in-differences analysis, SMI was associated with a statistically significantly lower proportion of households making any OOP expenditure on health in Honduras (Coef: -0.10, 95% CI: -0.20, 0.01), and insignificant elsewhere (Table 2). This is interpreted as the change in SMI areas being 10 percentage points lower than the change in comparison areas, when adjusted for covariates. Changes in the dollar amount of health expenditures were small in SMI areas of all countries, and SMI did not have a statistically significant impact on this outcome in any country in the difference-in-differences analysis (Supplementary Figure 1B, Table 2).

Impact of SMI on OOP prescription drug expenditure

The mean percent of total consumption spent on prescription drugs was similar between SMI intervention and comparison areas at baseline; in intervention areas, it ranged from 1.6% in Nicaragua to 2.6% in Chiapas, Mexico. Between baseline and follow-up, the mean percent of total consumption spent on prescription drugs decreased in the SMI areas of Guatemala (1.9% to 1.1%), Honduras (1.9% to 1.3%) and Nicaragua (1.6% to 1.3%) (Figure 2B). These represented substantial relative reductions, including a statistically significant 42% relative reduction in Guatemala. Simultaneously, the mean

percent of total consumption on prescription drugs increased in comparison areas in Honduras (1.2% to 2.3%) and Nicaragua (1.8% to 2.2%), while decreasing in Guatemala (2.0% to 1.2%). In Chiapas, the average proportion spent on prescription drugs was flat in both SMI (2.6% to 2.6%) and comparison (2.6% to 2.7%) areas. In the difference-in-differences analysis, SMI was associated with a statistically significant reduction in proportion of total household consumption spent on prescription drugs in Honduras (Coefficient (Coef): -0.02, 95% Confidence Interval (CI): -0.03, -0.00); the effect was insignificant in Guatemala (Coef: -0.01, 95% CI: -0.02, -0.00), Chiapas (Coef: 0.01, 95% CI: -0.01, 0.02) and Nicaragua (95% CI: -0.02, -0.04, 0.00) (Table 2).

The secondary outcome of percent of household making any prescription drug expenditure decreased in SMI areas of all four countries, while increasing in comparison areas of Honduras, Chiapas, and Nicaragua (Supplementary Figure 1C). The change was largest in Honduras, where the proportion of households making any prescription drug expenditure declined from 12.5% to 7.7% in SMI areas, while increasing from 9.1% to 18.1% in comparison areas (Supplementary Figure 1C). In the difference-in-differences analysis, SMI was associated with a statistically significant decrease in proportion of households making any prescription drug payment in Honduras (Coef: -0.11, 95% CI: -0.19, -0.02) (Table 2). SMI was not found to have a significant effect on the amount spent on prescription drugs in any country in the main analysis (Supplementary Figure 1D, Table 2).

When we restricted the analysis to households that visited a health facility in the past month (rather than all households), SMI was associated with a statistically significant reduction in proportion of total consumption on prescription drugs in Honduras (Coef: -0.02, 95% CI: -0.00, -0.01) and Guatemala (Coef: -0.05, 95% CI: -0.09, -0.00); making any OOP prescription drug expenditure in Honduras (Coef: -0.11, 95% CI: -0.22, -0.01) and Nicaragua (Coef: -0.13, 95% CI: -0.24, -0.02); and the amount of OOP expenditure on prescription drugs in Nicaragua (Coef: -4.58, 95% CI: -8.96, -0.21) (Supplementary Table 4).

SMI was associated with a statistically significant increase in the proportion of households with any non-prescription drug payment (Coef: 0.04, 95% CI: 0.01, 0.08) and the amount spent on non-prescription drugs in Guatemala (Coef: 0.24, 95% CI: 0.03, 0.46) (among the full population) (Supplementary Table 5). SMI was also associated with a decrease in the proportion of total consumption spent on non-prescription drugs in Chiapas among households visiting a health facility in the past month (Coef: -0.02, 95% CI: -0.03, -0.00). Elsewhere, the impact was insignificant.

Impact of SMI on care forgone due to cost

The proportion of households that reported forgoing care due to cost decreased in SMI areas of all countries, including large decreases in SMI areas of Guatemala (24.8% to 8.5%), Chiapas, Mexico (21.4% to 10.2%) and Nicaragua (25.5% to 13.1%) (Figure 2C). Comparison areas showed similar decreases as their respective SMI areas in Guatemala (26.2% to 5.7%) and Nicaragua (24.4% to 12.3%), and smaller decreases in Chiapas, Mexico (20.1% to 17.1%). In Honduras the proportion of households forgoing care due to cost changed minimally from baseline to follow-up in SMI areas (21.3% to 21.2%), but there was a large increase in Honduras comparison areas (19.1% to 33.0%). In the difference-in-differences analysis, SMI was significantly associated with less forgone care due to cost for SMI areas in Honduras

(Coef: -0.10, 95% CI: -0.20, -0.01), while the effects were smaller and insignificant in Chiapas, Nicaragua and Guatemala (Table 2). Additionally, in a sensitivity analyses that used a difference-in-differences framework but did not trim un-matched households or adjust for covariates or clustering of households within sampling units, SMI was associated with a significantly lower proportion of households forgoing care due to cost in both Honduras (Coef: -0.13, 95% CI: -0.18, -0.08) and Chiapas (Coef: -0.09, 95% CI: -0.13, -0.05).

Mediation analysis results

Shift from private to public sector

Most households reported a public facility as their usual facility at both baseline and follow-up. In SMI areas between 95% (Chiapas, Mexico) and 99% (Guatemala) of households said their usual facility was public at baseline (Figure 3A). With the exception of an increase in Guatemala comparison areas (+2 percentage points), the percent of households reporting a public facility as their usual facility declined in both SMI (Guatemala: -2; Honduras: -3; Chiapas: -4; Nicaragua: -3) and comparison (Honduras: -13; Chiapas: -4; Nicaragua: -6) areas of all countries. In the difference in differences analysis, we found that SMI was associated with a significant 'increase' (or lack of decrease) in having a public usual facility in Honduras (Coef: 0.09, 95% CI: 0.01, 0.16) (Table 3).

We observed consistent trends towards lower health expenditures in households who usually visited a public vs. private facility. For example, in Honduras, the mean percent of total expenditure on health was more than twice as high for households with a private usual facility (8.4%) than those with a public usual facility (3.4%); similar was true for mean percent of total consumption on prescription drugs (3.6% for private; 1.5% for public) (Figure 4A). Controlling for confounders (and pooling across intervention/control/baseline/follow-up), usually visiting a public facility was associated with significantly lower proportion of total consumption spent on health (Coef: -0.06, 95% CI: -0.12, -0.01) and proportion spent on prescription drugs (Coef: -0.03, 95% CI: -0.07, -0.00) in Honduras, but insignificant elsewhere.

We also observed that households usually visiting a public vs. private facility had a greater probability of having forgone care due to cost. For example, in Honduras, households with a public usual facility were more likely to have forgone care due to cost in the past year (23.0%) compared to those usually visiting a private facility (17.9%). However, there was not a statistically significant association between type of facility visited and care forgone due to cost in any country.

The formal mediation analysis only included those outcomes and countries for which SMI was determined to have had a significant impact in the difference-in-differences analysis, which included proportion total expenditure on prescription drugs, forgoing care due to cost, making any health expenditure, and making any prescription drug expenditure, all in Honduras only. We found that, in Honduras, usually visiting a public (vs. private) facility was a statistically significant mediator between SMI and spending a lower proportion of total consumption on prescription drugs (Coef: -0.01, 95% CI: -0.02, -0.004), and lower probability of making any OOP health expenditure (-0.01, 95% CI: -0.01, -0.00)

or any OOP prescription drug expenditure (-0.01, 95% CI: -0.01, -0.00) (Table 5). However, this mediator accounted for only 7.0% (proportion total consumption on prescription drugs), 4.3% (any health expenditure) and 4.1% (any prescription drug expenditure) of SMI's total impact. Usual facility type also had a statistically significant but inverse mediating effect on forgoing care due to cost in the past year (Coef: 0.01, 95% CI: 0.001, 0.01): households visiting public facilities were more likely to have forgone care due to cost (Figure 4A), whereas SMI reduced care forgone due to cost (Table 2).

Health status

The proportion of households where a woman or child had been sick in the past week declined in SMI areas of all countries. The largest reductions were in Honduras (51.7% to 41.6%) and Nicaragua (54.4% to 46.5%) (Figure 3B). This indicator also declined in all country's comparison areas except Honduras, where it remained the same. In the difference-in-differences analysis, SMI was associated with a statistically significant increase in having a woman or child sick in the past two weeks in Guatemala (due to large decreases in comparison areas relative to smaller decreases in intervention areas), and negative but statistically insignificant impacts in the other countries (Table 3).

We observed positive and statistically significant associations between having a woman or child sick in the past two weeks, and each of the primary financial protection outcomes (% total consumption on health; % total consumption on prescription drugs; care forgone due to cost) in all countries (Figure 4B, Table 4). For example, in Honduras, 29.2% of households with a sick woman or child in the past two weeks had forgone care due to cost in the past year, compared to 16.3% of households without a recently sick woman or child (Adjusted regression analysis coefficient: 0.14, 95% CI: 0.10, 0.18). Further, the percent of total consumption spent on health and prescription drugs was more than twice as much in Honduran households with a recently sick woman or child than households without (5.4% vs. 1.8% on total health; 2.2 vs. 0.9% on prescription drugs, respectively).

In the statistical mediation analysis of outcomes significantly impacted by SMI, recent illness of a woman or child in the household was found to be a statistically significant mediator of SMI's impact on proportion total expenditure on prescription drugs (Coef: -0.001, 95% CI: -0.002, -0.000), household forwent care due to cost (Coef: -0.01, 95% CI: -0.02, -0.004), household made any health expenditure (-0.01, 95% CI: -0.02, -0.01) and household made any prescription drug expenditure (-0.01, 95% CI: -0.01, -0.00) (these were all of the outcomes that SMI had significantly impacted) (Table 5). Recent woman/child illness accounted for 10.0%, 7.7%, 7.2% and 6.4% of SMI's impact on each of these outcomes, respectively.

Hospitalization rates

The proportion of households with a member hospitalized in the past year declined in both SMI intervention and comparison areas of all countries (Figure 3C). In the difference-in-differences analysis, SMI was associated with a statistically significant increase (or rather, a smaller decrease) in having a member hospitalized in Guatemala (Coef: 0.06, 95% CI: 0.02, 0.11); SMI did not have a statistically significant impact in the other countries.

There was a consistent (and in most cases statistically significant) positive association between having a household member hospitalized and higher (worse) financial protection outcomes (Table 4). For example, in Honduras, households with a member hospitalized spent a significantly greater proportion of total consumption on health (9.4% vs 2.7%; adjusted regression Coef: 0.06, 95% CI: 0.03, 0.08) and prescription drugs (3.0% vs. 1.4%), Coef: 0.02, 95% CI: 0.01, 0.03) and were (not statistically significantly) more likely to have forgone care due to cost (2.8% vs. 2.2%, Coef: 0.05, 95% CI -0.01, 0.11) (Figure 4C, Table 4).

In the statistical mediation analysis of SMI's significant impacts on financial protection (all in Honduras), hospitalization was not a statistically significant mediator for proportion total expenditure on prescription drugs, care forgone due to cost, making any health expenditure, and making any prescription drug expenditure.

Drug availability at public facilities

At baseline, the mean percent of tracer drugs available the day of the survey was similar in SMI and comparison areas, with overlapping confidence intervals (Figure 3D). From baseline to follow-up, the mean percent of tracer drugs observed to be available on the day of the health facility survey increased in SMI areas in Honduras (72% to 90%), Chiapas (57% to 84%) and Nicaragua (86% to 96%), and was flat in Guatemala (76% at both baseline and follow-up). Over the same period, the mean percent of tracer drugs available day-of at facilities in comparison areas increased in Honduras (70% to 87%) and Chiapas (55% to 64%) while declining in Guatemala (82% to 79%) and Nicaragua (92% to 88%). In the difference-in-differences analysis, SMI significantly increased the proportion of tracer drugs observed on the day of the survey in Chiapas (Coef: 0.20, 95% CI: 0.06, 0.33) and Nicaragua (Coef: 0.14, 95% CI: 0.04, 0.23) (Table 3). SMI did not have a significant impact on the proportion of drugs available in Guatemala nor Honduras. Results were consistent among facilities that had been sampled at both baseline and follow-up. The main drugs missing at follow-up in SMI areas included zinc (observed at 30% of ambulatory and 36% of basic/complete facilities), hydralazine (64% basic/complete) and diazepam (57% basic/complete) in Chiapas; iron (34% ambulatory, 40% basic/complete), ferrous sulfate (49% ambulatory, 100% basic/complete), hydralazine (60% basic/complete) and diazepam (22% basic/complete) in Guatemala; and magnesium sulfate (77% basic/complete) in Honduras (Supplementary Table 6). All other drugs were available in at least 80% of intervention ambulatory facilities in all countries at follow-up.

While our main measure of drug availability used a standard set of drugs across countries, as a sensitivity analysis we also examined a set of country-specific drugs tied to each country's payment indicators. For these indicators, SMI had a significant increase for the proportion of country-specific indicator drugs in stock day of the survey in Chiapas and Nicaragua; facilities having all country-specific indicator drugs in stock the day of the survey in Honduras, Chiapas and Nicaragua; and facilities having all country-specific indicator drugs in stock the day of the survey and three months prior without interruption in Honduras, Chiapas and Nicaragua (Supplementary Table 7A).

In descriptive analysis, there were mixed associations between the availability of drugs in a facility and the primary financial protection outcomes at households visiting that facility in the past month (Figure 4D), including when split by ambulatory vs. basic/complete EONC level (Supplementary Figure 3).

Inconsistent associations were also observed for the secondary financial protection outcomes related to prescription drug expenditures including whether any prescription drug expenditure was made, and the amount spent on prescription drugs (Supplementary Figure 4). The confidence intervals were large and overlapping, due in part to the reduced sample size of households that could linked to a recently visited facility with a survey. When controlling for covariates, we found no significant associations between the proportion of tracer drugs available at the facility the day of the survey and the financial protection outcomes, including the proportion of household income spend on prescription drugs (Guatemala - Coef: -0.05, 95% CI: -0.13, 0.02; Honduras – Coef: 0.01, 95% CI: -0.05, 0.06; Chiapas – Coef: 0.01, 95% CI: -0.04, 0.06; Nicaragua – Coef: 0.03, 95% CI: -0.04, 0.09) (Table 4). Among the secondary indicators related to prescription drug expenditures, visiting a facility with more tracer drugs observed in stock was associated with a reduction in households making any prescription drug expenditure in Guatemala (Coef: -0.27, 95% CI: -0.50, 0.03), but otherwise there was no significant association with whether the household made any OOP expenditure on prescription drugs, nor the dollar amount of OOP prescription drug expenditures. Associations were also inconsistent and insignificant when considering availability of country-specific indicator drugs rather than the standard set of tracers, with the exception of a significant association between continuous availability of country-specific drugs for three months and proportion of total consumption spent on prescription drugs in Honduras (Supplementary table 7B, 7C).

In the formal mediation analysis, we found no indications that the proportion of drugs available at a facility had a mediating effect on the financial protection outcomes (Table 5). However, we also noted that among this smaller sub-set of households that could be linked to a recently visited health facility, the overall impacts of SMI on the financial protection outcomes were also smaller in magnitude and statistically insignificant, as compared to among the full sample.

DISCUSSION

We find evidence that SMI significantly improved several financial protection indicators in the poorest communities of Honduras, including proportion of total consumption spent on prescription drugs, households forgoing care due to cost, households making any OOP health expenditure, and households making any OOP prescription drug expenditure. SMI did not have a significant impact on any financial protection outcomes in Guatemala, Nicaragua and Chiapas (Mexico) among the full study population; it did improve select prescription drug expenditure outcomes among households visiting a facility in the past month in Guatemala and Nicaragua (in addition to Honduras). We find no indications that SMI negatively affected our main financial protection outcomes in the intervention areas of any country, though some did worsen in comparison areas. The significant impacts in Honduras appear to be mediated in part by a decreased incidence of illness (or perceived illness) among women and children in SMI areas, though these mediators account for a small proportion of the total impact (<10% each). Notably, Honduras was the only country with fully randomized assignment of municipalities to intervention and control, increasing confidence in the findings, and there are no indications of a substitution effects from non-prescription drug expenditures in Honduras.

Given that the world is not on track to achieve the financial protection Sustainable Development Goal, new approaches to improving financial protection are needed [19]. Existing literature is largely focused on broad policy reforms, risk-pooling mechanisms and financially-oriented interventions such as loan and microcredit programs, with a bias to high-income settings [20–22]. In contrast, our study asks whether financial protection was indirectly influenced by the first and second phases of the SMI intervention, which took a broader approach to improving the supply and delivery of services in poor communities of middle-income countries. Null results in several countries indicate that the SMI approach does not consistently lead to financial protection improvements, nor does it worsen financial burden on households (a potential concern with increased utilization). However, the significant improvements among all households in Honduras - and households with a recent health facility visit in Guatemala and Nicaragua - suggest that it may be feasible under certain conditions to improve financial protection, particularly prescription drug expenditures, through improving overall service availability and quality. In some cases, SMI's impacts in Honduras may even be comparable to impacts observed for more traditional financial protection interventions. For example, SMI was associated with a (non-significant) two percentage-point reduction in the proportion of total consumption spent on health in Honduras, while a recent pilot study of Ethiopia's insurance system found it was associated with a significant 1 percentage-point reduction in the same outcome, as compared to non-pilot districts [23].

The mixed results across countries could reflect several things, including differing ambitions and strategic scope going into the initiative, contextual factors needed for success, and quality of implementation. For example, in Guatemala, SMI was associated with worsening outcomes for women/children illness, hospitalization rates and non-prescription drug expenditures; Guatemala was also the only country where SMI did not significant improve availability of either the standard tracer or country-specific drugs. This aligns with field reports that Guatemala experienced multiple implementation challenges, including drug acquisition problem during the second phase of SMI. A qualitative process evaluation is needed to understand why SMI has been more successful in some countries than others and if there are certain contextual pre-conditions for success. We also note that in some cases (e.g.: fewer household forgoing care due to cost in Honduras and improved drug availability in Nicaragua) the positive difference-in-differences effect of SMI was driven in part by worsening outcomes in non-SMI areas. Process evaluations should also investigate the possibility that resources were diverted to SMI areas from comparison areas, which would be a negative unintended consequence of the incentives established by the initiative, and what the implications of this would be if this type of model were implemented at scale (e.g.: it may still be ethical to redistribute some resources and attention from higher income to lower income parts of the country).

In Honduras, reduced proportion of total consumption on prescription drugs, fewer households making any health or prescription drug payments and less care forgone due to cost appeared to be tied, in part, to overall improvements in health, or at least self-perceived health. This would align with the SMI theory of change, which posits that improving the availability and quality of care - especially for preventive services - can improve population health [24]. Interestingly, despite reportedly lower rates of woman/child illness, the overall proportion of households visiting a facility in the past month was very similar across intervention and comparison, and baseline and follow-up. This would suggest that while the incidence of visits has stayed the same, the balance has shifted to more preventive and fewer

curative visits in SMI areas, which would be consistent with lower expenditures, especially on prescription drugs. Indeed, SMI was associated with a (non-statistically significant) decrease in hospitalization rates in Honduras. Additionally, an impact evaluation has shown early indications of SMI significantly improving the quality and timeliness of antenatal care in Honduras [Source: Aim 3]. Other preventive services, such as vaccination rates, appear to be improving as well, though the causal impact of SMI on these is yet to be evaluated [25, 26].

SMI's positive financial protection impacts in Honduras also appear to be driven by relative utilization of the public vs private sector. Across intervention and comparison areas in all countries, there was an overall trend towards increased utilization of the private sector. However, in Honduras, this shift was much more pronounced in the comparison areas compared to the intervention areas. A quantitative baseline study of SMI populations in Guatemala, Honduras and Nicaragua demonstrated that women take facility capacity into account when choosing a delivery facility [11]. This was further reinforced by qualitative work in the SMI regions of Chiapas, showing that perceptions of low quality care at public facilities can drive women to the private sector [12, 27]. Therefore, one interpretation is that by improving supply availability and service quality at public facilities, SMI attracted women to seek free/low-cost public services rather than paying in the private sector. Another more concerning possibility, given the large shift to private sector in the comparison areas of Honduras, is that SMI incentivized Honduras to divert public facility resources away from comparison area and towards the intervention areas. While this could theoretically result in lower quality and increased utilization of the private sector in comparison areas, this is refuted by the fact that public facility drug availability improved in Honduras' comparison areas as well as intervention areas.

The mediation analysis refuted our hypothesis that improved drug availability mediated SMI's impact on financial protection. This was surprising, given that SMI affected the prescription drug expenditure outcomes more than the total expenditure outcomes. Previous studies have found a relationship between stock-outs and OOP expenditure for HIV and malaria drugs in African settings, including a study in Tanzania finding that malaria drug stock outs increased household expenditure by 21% in malaria patients [7, 28]. It is possible that our study was underpowered to detect an effect, or that our drug availability indicator was poorly specified (e.g. it treated all drugs equally, whereas some drugs may drive costs more than others). However, we also note other contradicting evidence: for instance, SMI reduced the proportion of total expenditure on prescription drugs among households recently visiting a facility in Guatemala without improving any of the drug availability measures; while in Chiapas, SMI increased availability of the country-specific indicator drugs, but did not improve any OOP prescription drug expenditure outcomes.

While we did not find it to be part of the causal pathway for financial protection, the increase in drug availability in all four countries (including statistically significant increases over the comparison areas in Chiapas and Nicaragua for the standard tracer drugs, and in Honduras, Chiapas and Nicaragua for country-specific drugs) is still notable, and confirms early indications of reduced stock-outs following SMI's first phase [29]. Inconsistent drug supply is a known barrier to high quality services and utilization in many low and middle-income countries [30]. Yet, few interventions have been shown to improve availability in rigorous studies and the effectiveness of specific interventions varies by context [31, 32].

The impact of SMI is comparable or larger than other known interventions. For instance, an RCT of a supervisory touring program in Zimbabwe found a 13% improvement in the availability of 15 essential indicator items in the intervention vs control group, while the SMI model was associated with a 20% improvement in Chiapas and a 14% improvement in Nicaragua for the standard tracer drugs [33]. The SMI model offered incentives and resources for countries to design and implement solutions suited to their own system and barriers, rather than intervening with a uniform solution. Each country was pursuing financial incentives tied to improved stock, but the specific interventions to achieve this varied by country, including strengthening the Logistical Management Unit to ensure drug availability in Chiapas and providing incentives to health facilities for maintaining adequate stock in Honduras. Perhaps not surprisingly, SMI had a larger impact on the availability of drugs that mattered for their country-specific payment-linked performance indicators, demonstrating country responsiveness to these incentives.

The findings are subject to important limitations. First, there is the fact that intervention and comparison areas were not randomly assigned in Guatemala, Chiapas and Nicaragua, though efforts were made to match municipalities on sociodemographic characteristics; however, assignment was random in Honduras, the country for which we find the most significant SMI impacts. We also note the possibility for spill-over occurring between intervention and comparison areas, given that some SMI interventions (e.g.: norms and policy changes) were designed to be integrated with the broader health system and could not be strictly contained to SMI areas – this could bias towards a null result and lead us to under-estimate the impact of SMI. Another source of bias towards a null result is if patients from comparison areas were attracted to attend facilities in SMI areas when they saw that the availability and quality of supplies and services was improving. In addition, the sample sizes were not generated specifically for this analysis, and therefore the study was likely underpowered to answer some questions, particularly among the smaller sample of linked health facilities and households. Conversely, we did not account for multiple comparisons, meaning we may be over-estimating impact. There are also limitations related to the measurement of expenditure outcomes, as the household expenditure data are based on recall, which can introduce bias and measurement error [13]. We mitigated this limitation by using consistent questionnaires across countries and survey rounds, so that, at minimum, any bias remained consistent and could theoretically be addressed by the difference in differences framework. Further, we did not account for insurance reimbursements in the calculation of prescription drug expenditures (though we did account for them for the secondary CHE and total health expenditure outcomes), as we did not know if the reimbursements were for prescription drugs. However, the frequency and amount of these reimbursements was extremely low. Finally, the accuracy and completeness of history stock-availability records for the country-specific indicator-linked drugs may have been affected by the initiative, so we did not consider historic stock-availability as a primary outcome.

CONCLUSION

SMI significantly improved several financial protection outcomes in Honduras; certain outcomes were also improved among the sub-set of households visiting a health facility in the past month in Honduras,

Guatemala and Nicaragua. The significant impacts in Honduras appear to be mediated by reduced rates of self-reported woman/child illness in SMI intervention areas relative to comparison areas. SMI also significantly improved drug availability in several countries, but this was not found to be part of the causal pathway between SMI and financial protection. Additional research is needed to understand the contextual factors determining why SMI improved financial protection in some settings but not others.

ACKNOWLEDGEMENTS

We thank the women and health care workers who gave their time and assistance to make this study possible. We also thank the many individuals involved in collecting the data including El Colegio de la Frontera Sur-Mexico, The Centre for Research and Health Studies at the National Autonomous University of Nicaragua (CIES-UNAN), UNIMER, and Fundación FES.

FIGURES

TABLES

CITATIONS

1. Wagstaff A, Flores G, Hsu J, Smitz M-F, Chepynoga K, Buisman LR, et al. Progress on catastrophic health spending in 133 countries: a retrospective observational study. Lancet Glob Health. 2018;6:e169–79.

2. Knaul FM, Wong R, Arreola-Ornelas H, Méndez O, Network on Health Financing and Social Protection in Latin America and the Caribbean (LANET). Household catastrophic health expenditures: a comparative analysis of twelve Latin American and Caribbean Countries. Salud Publica Mex. 2011;53 Suppl 2:s85-95.

3. PAHO. Strategy for Universal Access to Health and Universal Health Coverage. 2014. http://www.paho.org/uhexchange/index.php/en/uhexchange-documents/technical-information/26-strategy-for-universal-access-to-health-and-universal-health-coverage/file.

4. AMRO/PAHO. Health Financing and Financial Protection in the Americas - Working Paper. 2017. http://www.who.int/health_financing/events/who_paho_uhc_day_2017_report_web.pdf.

5. Mokdad AH, Colson KE, Zúñiga-Brenes P, Ríos-Zertuche D, Palmisano EB, Alfaro-Porras E, et al. Salud Mesoamérica 2015 Initiative: design, implementation, and baseline findings. Popul Health Metr. 2015;13:3.

6. Saksena P, Xu K, Durairaj V. The drivers of catastrophic expenditure: outpatient services, hospitalization or medicines? Background paper. World Health Organization; 2010. http://www.who.int/healthsystems/topics/financing/healthreport/21cataclearance.pdf.

7. Baglione Q. Stock outs cause catastrophic expenditure and further impoverishment of ART patients in South Africa. Paris, France; 2017.

8. Moye-Holz D, van Dijk JP, Reijneveld SA, Hogerzeil HV. Policy approaches to improve availability and affordability of medicines in Mexico – an example of a middle income country. Glob Health. 2017;13:53.

9. Anson A, Ramay B, de Esparza AR, Bero L. Availability, prices and affordability of the World Health Organization's essential medicines for children in Guatemala. Glob Health. 2012;8:22.

10. Emmerick ICM, Luiza VL, Camacho LAB, Vialle-Valentin C, Ross-Degnan D. Barriers in household access to medicines for chronic conditions in three Latin American countries. Int J Equity Health. 2015;14. doi:10.1186/s12939-015-0254-z.

11. Hernandez B, Colombara DV, Gagnier MC, Desai SS, Haakenstad A, Johanns C, et al. Barriers and facilitators for institutional delivery among poor Mesoamerican women: a cross-sectional study. Health Policy Plan. 2017;32:769–80.

12. Dansereau E, Schaefer A, Hernández B, Nelson J, Palmisano E, Ríos-Zertuche D, et al. Perceptions of and barriers to family planning services in the poorest regions of Chiapas, Mexico: a qualitative study of men, women, and adolescents. Reprod Health. 2017;14. doi:10.1186/s12978-017-0392-4.

13. Lavado RF, Brooks BP, Hanlon M. Estimating health expenditure shares from household surveys. Bull World Health Organ. 2013;91:519-524C.

14. PPP conversion factor, GDP (LCU per international $) | Data. https://data.worldbank.org/indicator/PA.NUS.PPP. Accessed 20 Nov 2018.

15. CPI Inflation Calculator. https://data.bls.gov/cgi-bin/cpicalc.pl. Accessed 20 Nov 2018.

16. Iacus SM, King G, Porro G. Causal Inference Without Balance Checking: Coarsened Exact Matching. Polit Anal. 2012;20:1--24.

17. Fairchild AJ, McDaniel HL. Best (but oft-forgotten) practices: mediation analysis12. Am J Clin Nutr. 2017;105:1259–71.

18. mediation: R Package for Causal Mediation Analysis | Tingley | Journal of Statistical Software. doi:10.18637/jss.v059.i05.

19. Verguet S, Woldemariam AT, Durrett WN, Norheim OF, Kruk ME. Is the sustainable development goal target for financial risk protection in health realistic? BMJ Glob Health. 2017;2. doi:10.1136/bmjgh-2016-000216.

20. Essue BM, Kimman M, Svenstrup N, Lindevig Kjoege K, Lea Laba T, Hackett ML, et al. The effectiveness of interventions to reduce the household economic burden of illness and injury: a systematic review. Bull World Health Organ. 2015;93:102-112B.

21. Ekman B. Community-based health insurance in low-income countries: a systematic review of the evidence. Health Policy Plan. 2004;19:249–70.

22. Gertler P, Levine DI, Moretti E. Do microfinance programs help families insure consumption against illness? Health Econ. 2009;18:257–73.

23. Mebratie AD, Sparrow R, Yilma Z, Abebaw D, Alemu G, Bedi AS. The impact of Ethiopia's pilot community based health insurance scheme on healthcare utilization and cost of care. Soc Sci Med. 2019;220:112–9.

24. Regalia F, Eichler R, Gigli S, Cleason M, Tapia Conyer R, Hernandez G, et al. Not Your Cookie-Cutter Results-Based Aid Initiative: Salud Mesoamerica Initiative's Experience Improving Health for the Poorest in Mesoamerica. Technical Notes. Inter-American Development Bank; 2017. doi:10.18235/0000855.

25. Ozawa S, Mirelman A, Stack ML, Walker DG, Levine OS. Cost-effectiveness and economic benefits of vaccines in low- and middle-income countries: A systematic review. Vaccine. 2012;31:96–108.

26. IHME. SMI-Honduras Household Census and Survey Data Quality Report, Second Follow-up Measurement. 2018.

27. El Bcheraoui C, Palmisano E, Dansereau E, Schaefer A, Woldeab A, Hernandez B, et al. Salud Mesoamérica Initiative Process Evaluation: Final Report. Institute for Health Metrics and Evaluation (IHME); 2017.

28. Mikkelsen-Lopez I, Tediosi F, Abdallah G, Njozi M, Amuri B, Khatib R, et al. Beyond antimalarial stock-outs: implications of health provider compliance on out-of-pocket expenditure during care-seeking for fever in South East Tanzania. BMC Health Serv Res. 2013;13:444.

29. Mokdad A. Supply-side interventions to improve health: findings from the Salud Mesoamérica Initiative. PLOS ONE. Forthcoming.

30. Mariko M. Quality of care and the demand for health services in Bamako, Mali: the specific roles of structural, process, and outcome components. Soc Sci Med. 2003;56:1183–96.

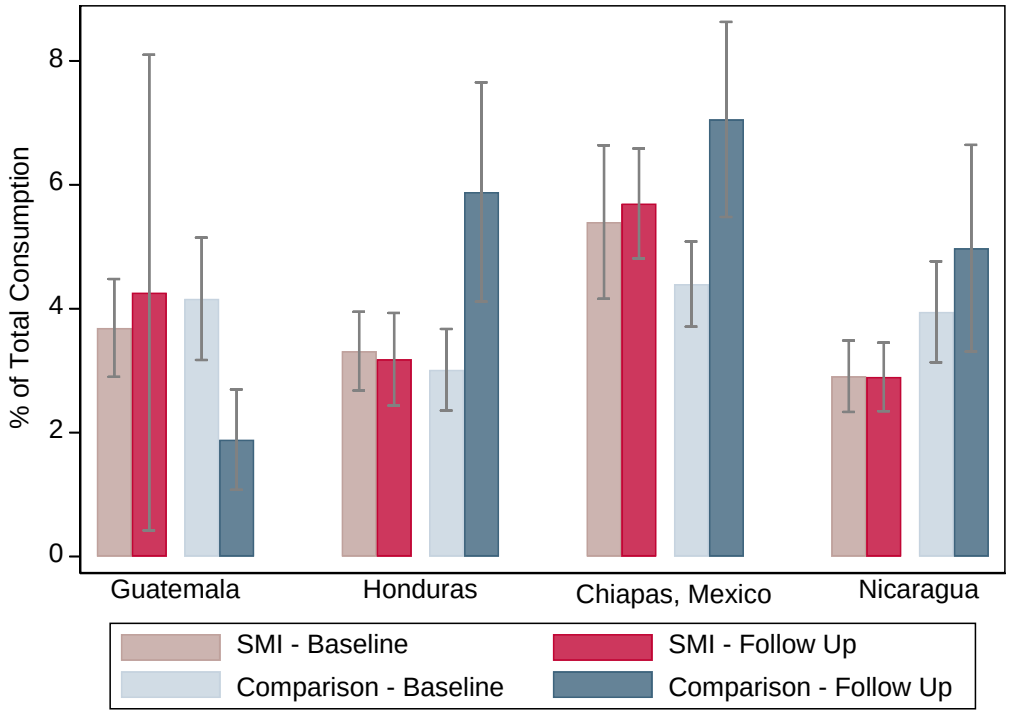
31. Yadav P. Health Product Supply Chains in Developing Countries: Diagnosis of the Root Causes of Underperformance and an Agenda for Reform. Health Syst Reform. 2015;1:142–54.

32. Nunan M, Duke T. Effectiveness of pharmacy interventions in improving availability of essential medicines at the primary healthcare level. Trop Med Int Health. 2011;16:647–58.

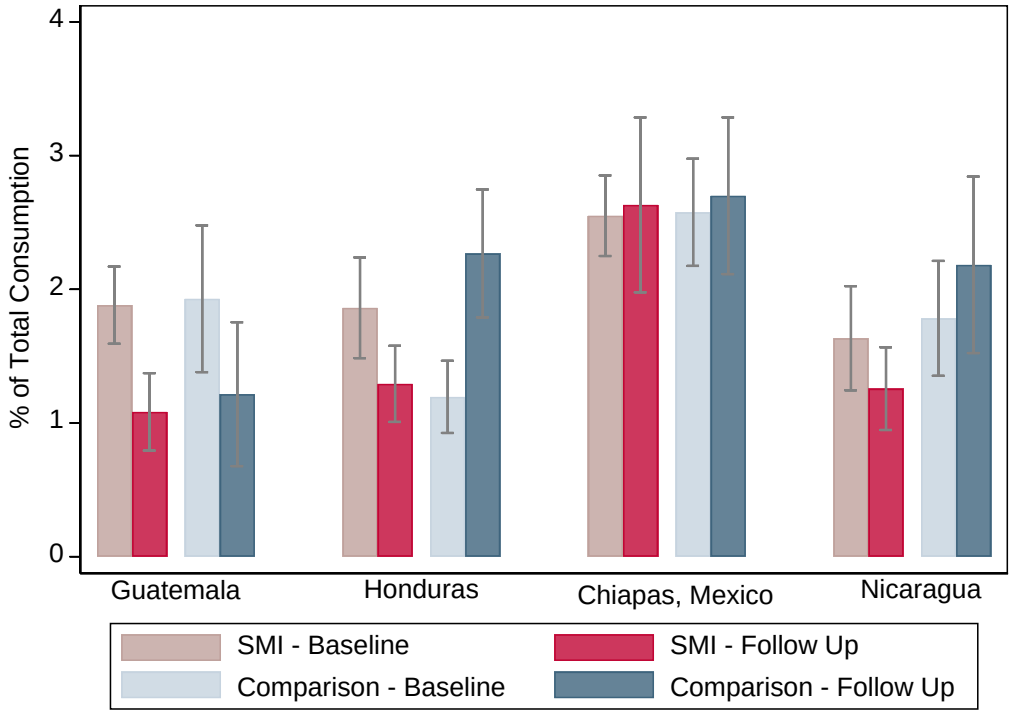
33. Trap B, Todd CH, Moore H, Laing R. The impact of supervision on stock management and adherence to treatment guidelines: a randomized controlled trial. Health Policy Plan. 2001;16:273–80.

Figure 2: Financial protection outcomes in SMI intervention and comparison areas at baseline and follow-up, among matched households

A: Mean percent of total household consumption spent on health, minus insurance reimbursements



B: Mean percent of total household consumption spent on prescription drugs



C: Proportion of households that reported forgoing care due to cost in the past year

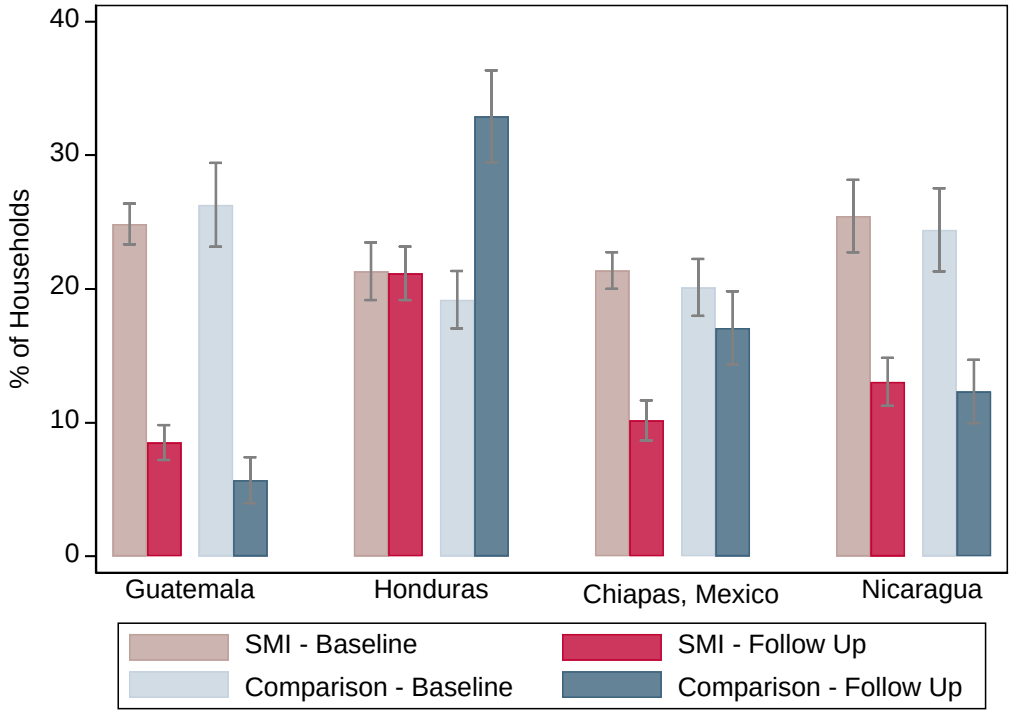
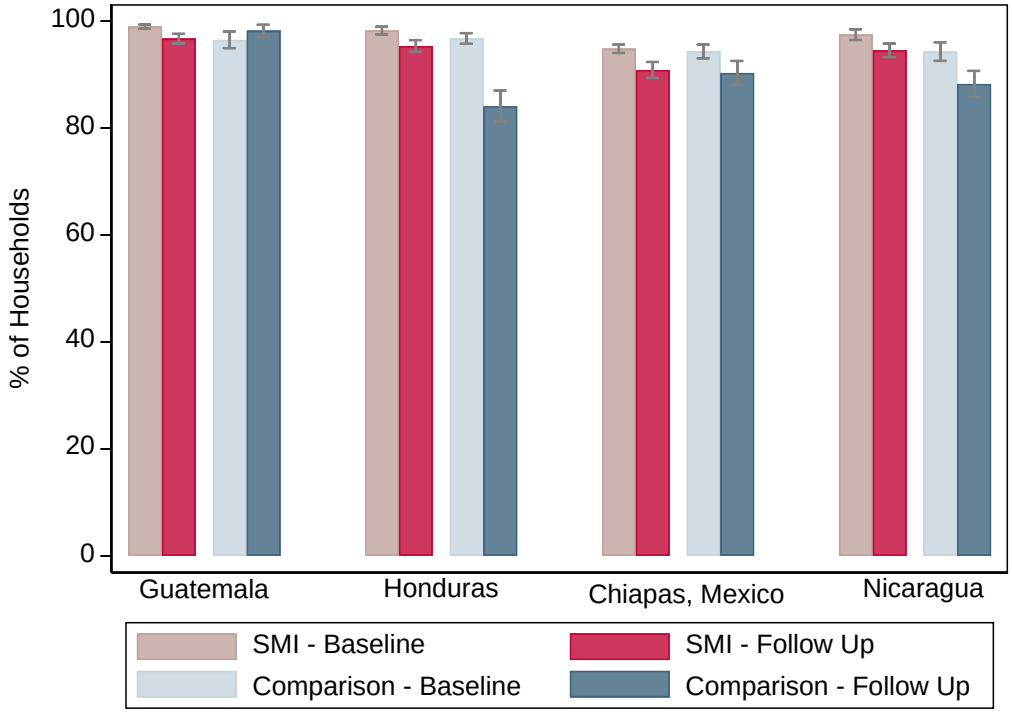
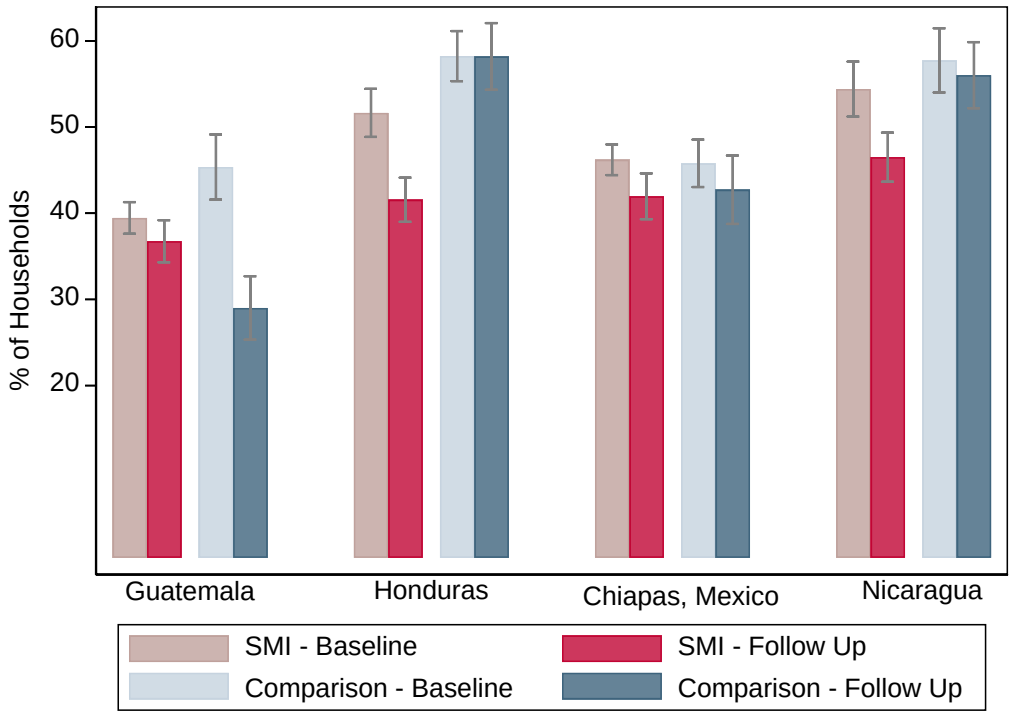


Figure 3: Change in factors hypothesized to mediate the impact of SMI on financial protection outcomes in SMI intervention and comparison areas at baseline and follow-up, among matched households

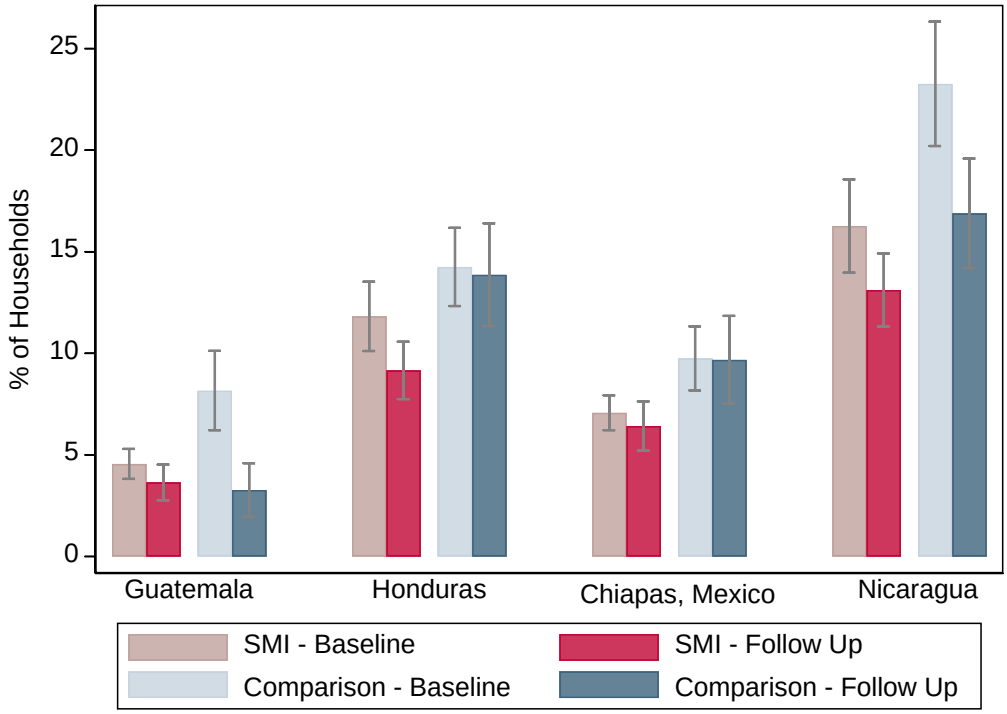
A: Percent of households reporting a public facility as their usual facility



B: Percent of households where a woman or child has been sick in the past 2 weeks



C: Percent of households where a member was hospitalized in the past year



D: Mean percent of tracer drugs observed to be in stock on the day of the health facility survey

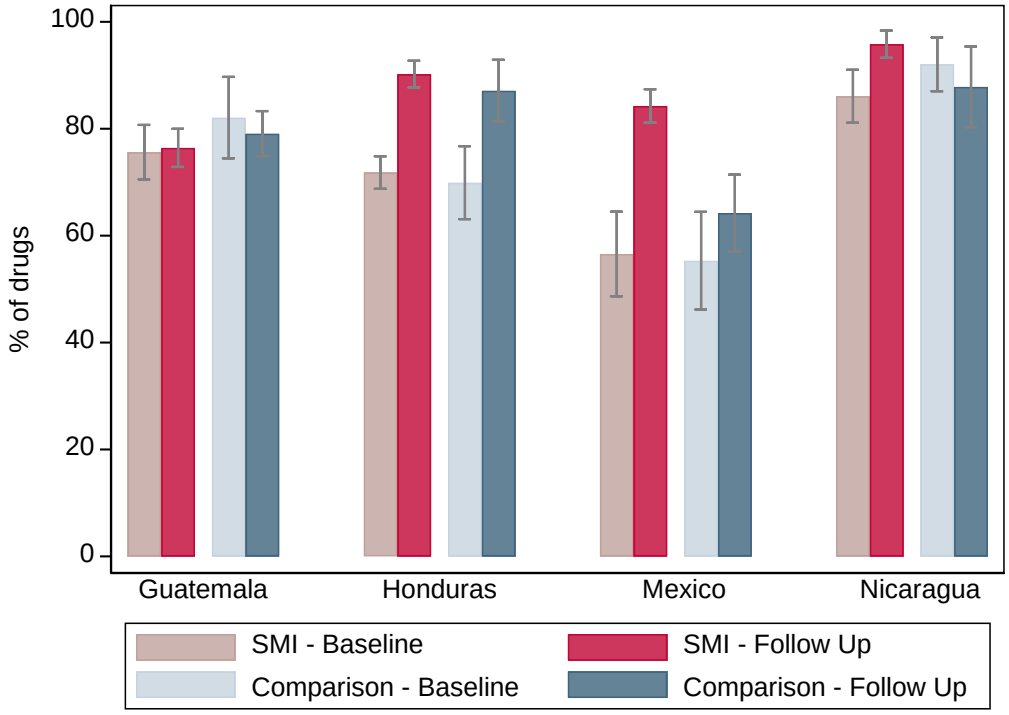
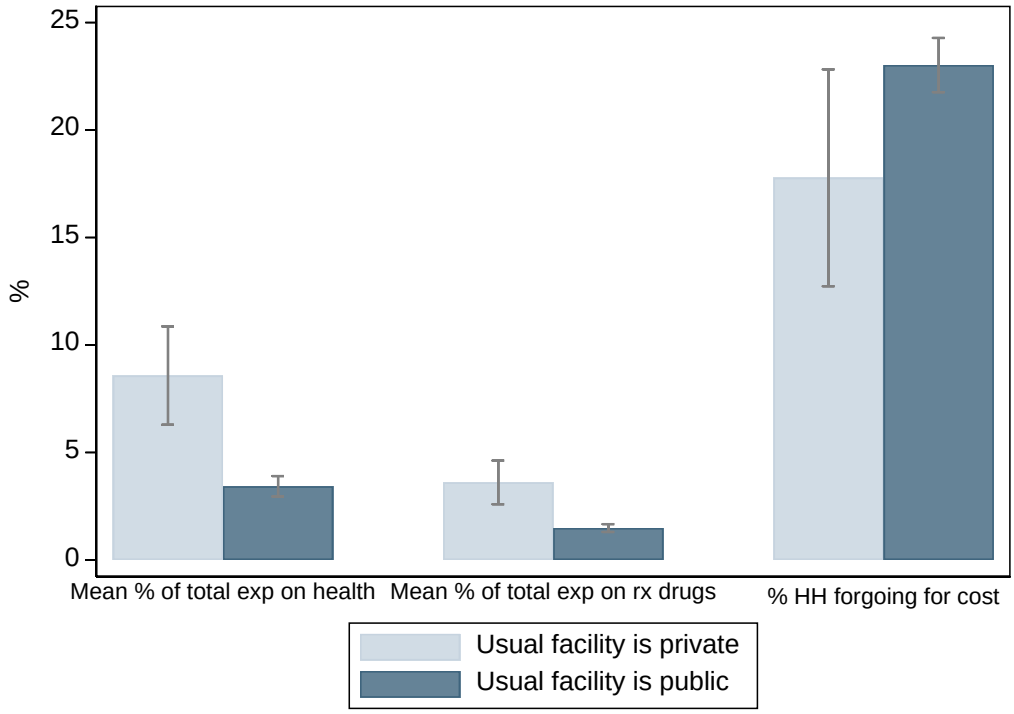
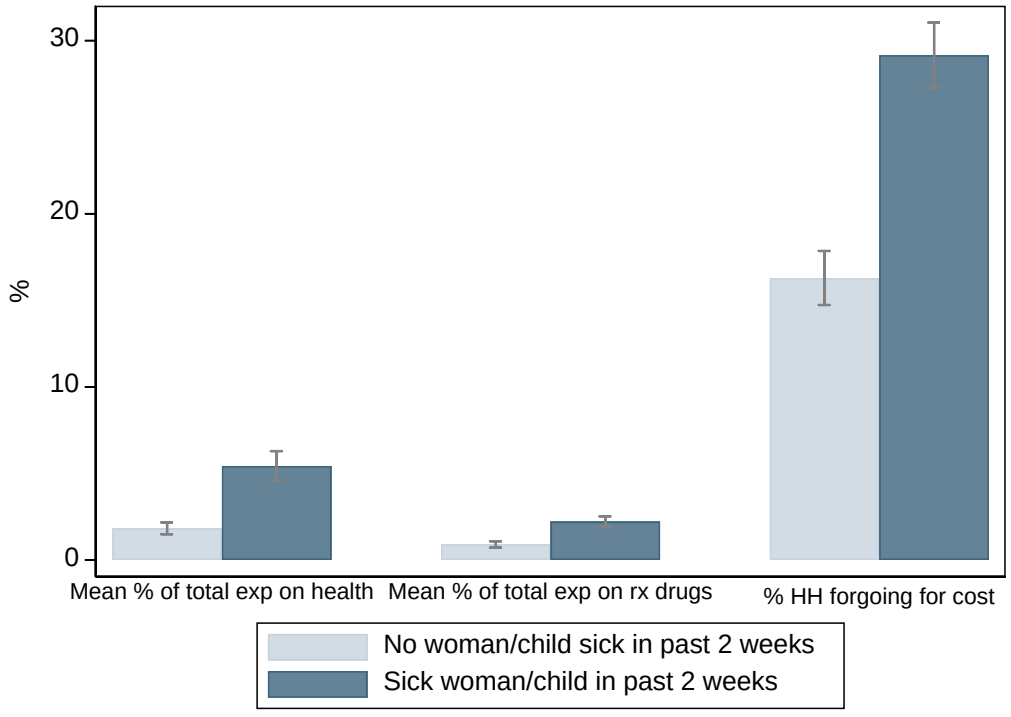


Figure 4: Associations between hypothesized mediators and financial protection indicators in Honduras, pooled across baseline/follow-up and intervention/comparison

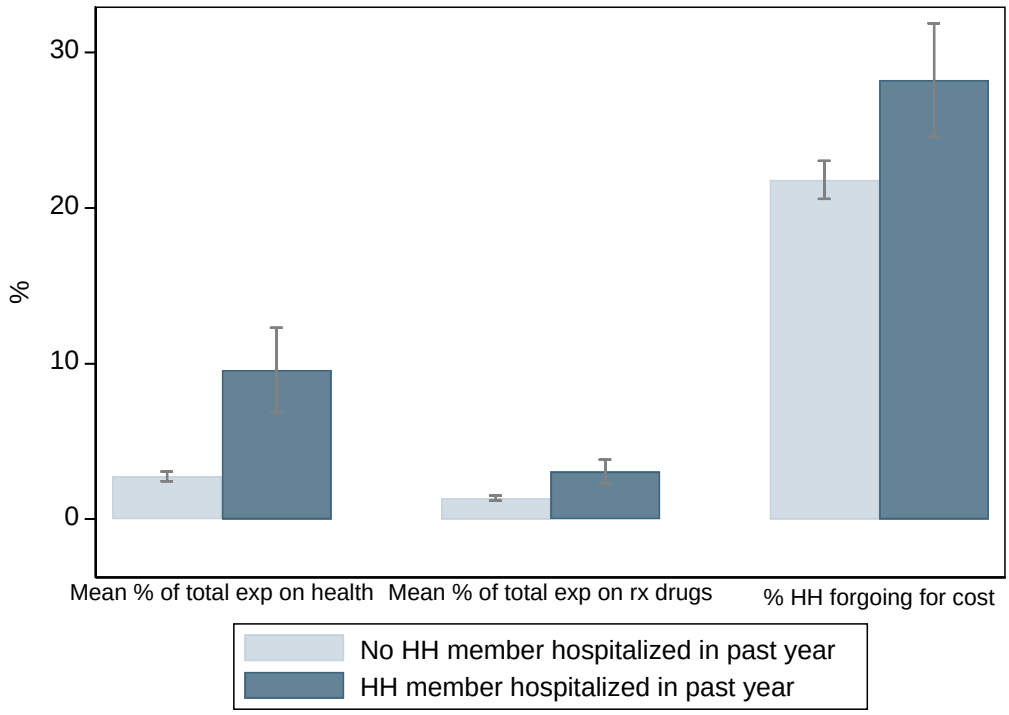
A: Financial protection indicators, by whether the household reported a public facility as their usual facility



B: Financial protection indicators, by whether any woman or child in the household was sick in the past 2 weeks



C: Financial protection indicators, by whether the household reported having a member hospitalized in the past year



D: Mean percent of total household expenditure spent on health (in total) and on prescription drugs, by percent of tracer drugs observed to be in stock at the facility recently visited by that household

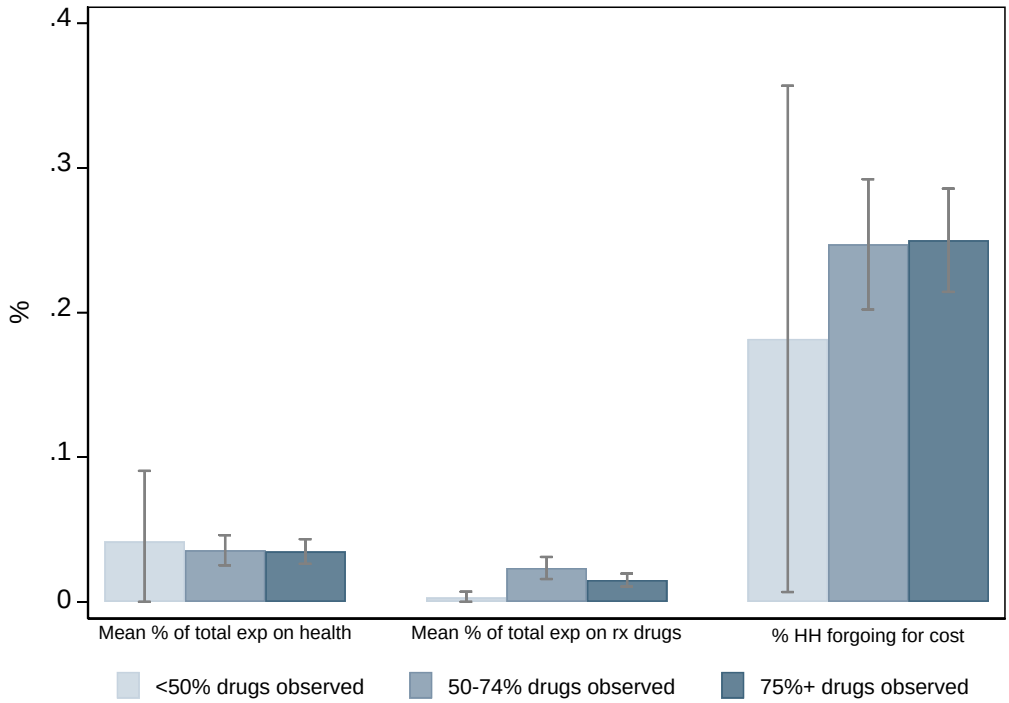


Table 1: Sample of households at baseline and follow-up in SMI intervention and comparison areas, after matching

	SMI intervention areas		Comparison areas		
	Baseline	Follow Up	Baseline	Follow Up	Total
Guatemala	3,079	1,760	763	703	6,305
Honduras	1,393	1,609	1,295	722	5,019
Chiapas, Mexico	3,444	1,576	1,367	727	7,114
Nicaragua	996	1,358	737	746	3,837
Total	8,912	6,303	4,162	2,898	22,275

Table 2: Estimated impact of SMI on household out of pocket prescription drug expenditure indicators in difference-in-differences analyses, following matching and controlling for covariates

Among all households	Guatemala	Honduras	Chiapas, Mexico	Nicaragua
Primary outcome				
Proportion of total household consumption spent on health	-0.00 [-0.02, 0.02]	-0.02 [-0.05, 0.01]	-0.01 [-0.04, 0.02]	-0.03 [-0.08, 0.01]
Proportion of total household consumption spent on prescription drugs	-0.01 [-0.02, 0.00]	-0.02* [-0.03, -0.00]	0.01 [-0.01, 0.02]	-0.02 [-0.04, 0.00]
Household forwent care due to cost in the past year	0.03 [-0.04, 0.11]	-0.10* [-0.20, -0.01]	-0.06 [-0.13, 0.02]	-0.02 [-0.10, 0.06]
Secondary outcomes				
Household made any OOP expenditure on health	0.06 [-0.00, 0.12]	-0.10* [-0.20, -0.01]	0.02 [-0.05, 0.10]	-0.06 [-0.16, 0.03]
OOP expenditure on health (2018 USD PPP)	0.28 [-1.18, 1.75]	-0.46 [-7.13, 6.21]	-0.67 [-5.05, 3.71]	-4.13 [-11.68, 3.42]
Household made any OOP expenditure on prescription drugs	-0.00 [-0.05, 0.04]	-0.11* [-0.19, -0.02]	-0.00 [-0.06, 0.06]	-0.05 [-0.14, 0.04]
OOP expenditure on prescription drugs (2018 USD PPP)	-0.57 [-1.61, 0.46]	-2.27 [-4.84, 0.30]	0.70 [-0.69, 2.09]	-2.80 [-6.87, -1.26]
N	6305	5019	7114	3837

* p<0.05, ** p<0.01, *** p<0.001

Table 3: Estimated impact of SMI on factors hypothesized to mediate its impact on financial protection outcomes, from a difference-in-differences analysis

	Guatemala	Honduras	Chiapas, Mexico	Nicaragua
Usual health facility is public	-0.01 [-0.04, 0.02]	0.09* [0.01, 0.16]	0.02 [-0.04, 0.08]	0.10 [-0.01, 0.22]
Woman or child in household sick in past two weeks	0.13* [0.02, 0.24]	-0.07 [-0.17, 0.03]	-0.04 [-0.16, 0.08]	-0.08 [-0.19, 0.03]
Household member hospitalized in past year	0.06** [0.02, 0.11]	-0.04 [-0.10, 0.01]	0.00 [-0.05, 0.05]	0.03 [-0.06, 0.12]
Proportion of tracer drugs in stock day of the survey	0.04 [-0.07, 0.15]	0.02 [-0.06, 0.10]	0.20** [0.06, 0.33]	0.14** [0.04, 0.23]

* p<0.05, ** p<0.01, ***

Table 4: Association between hypothesized mediators and financial protection outcomes in adjusted regression analysis				
Association between normally visiting a public facility and financial protection outcomes	Guatemala	Honduras	Chiapas, Mexico	Nicaragua
Association with proportion of total consumption spent on health	-0.04 [-0.08, 0.01]	-0.06* [-0.12, -0.01]	-0.03 [-0.08, 0.03]	-0.02 [-0.06, 0.02]
Association with proportion of total consumption spent on prescription drugs	-0.01 [-0.04, 0.02]	-0.03* [-0.07, -0.00]	-0.01 [-0.03, 0.00]	-0.02 [-0.04, 0.01]
Association with whether the household forwent care in the past year due to cost	0.00 [-0.11, 0.11]	-0.04 [-0.15, 0.07]	-0.03 [-0.09, 0.03]	0.03 [-0.05, 0.10]
N	6305	5019	7114	3837
Association between woman/child illness in the past 2 weeks and financial protection outcomes	Guatemala	Honduras	Chiapas, Mexico	Nicaragua
Association with proportion of total consumption spent on health	0.02* [0.00, 0.04]	0.04*** [0.03, 0.05]	0.04*** [0.03, 0.05]	0.04*** [0.02, 0.05]
Association with proportion of total consumption spent on prescription drugs	0.02*** [0.01, 0.02]	0.02*** [0.01, 0.02]	0.02*** [0.01, 0.02]	0.02*** [0.01, 0.02]
Association with whether the household forwent care in the past year due to cost	0.14*** [0.10, 0.18]	0.14*** [0.10, 0.18]	0.06*** [0.03, 0.09]	0.14*** [0.10, 0.18]
N	6305	5019	7114	3837
Association between hospitalization in the past year and financial protection outcomes	Guatemala	Honduras	Chiapas, Mexico	Nicaragua
Association with proportion of total consumption spent on health	0.22* [0.01, 0.44]	0.06*** [0.03, 0.08]	0.08*** [0.04, 0.13]	0.04*** [0.01, 0.07]
Association with proportion of total consumption spent on prescription drugs	0.05** [0.02, 0.08]	0.02** [0.01, 0.03]	0.02** [0.01, 0.03]	0.02* [0.00, 0.03]
Association with whether the household forwent care in the past year due to cost	0.20*** [0.11, 0.29]	0.05 [-0.01, 0.11]	0.10*** [0.04, 0.16]	0.11*** [0.07, 0.16]
N	6305	5019	7114	3837

* p<0.05, ** p<0.01, *** p<0.001

(cnt) Table 4: Association between hypothesized mediators and financial protection outcomes in adjusted regression analysis

Association between proportion of drugs availability at recently visited facility and financial protection outcomes	Guatemala	Honduras	Chiapas, Mexico	Nicaragua
Association with proportion of total consumption spent on health	-0.08 [-0.16, 0.01]	0.02 [-0.07, 0.11]	-0.04 [-0.10, 0.03]	0.09 [-0.05, 0.22]
Association with proportion of total consumption spent on prescription drugs	-0.05 [-0.13, 0.02]	0.01 [-0.05, 0.06]	0.01 [-0.04, 0.06]	0.03 [-0.04, 0.09]
Association with whether the household forwent care in the past year due to cost	0.09 [-0-.24, 0.43]	0.13 [-0.13, 0.39]	-0.15 [-0.32, 0.02]	-0.09 [-0.48, 0.30]
N	533	946	413	388

* p<0.05, ** p<0.01, *** p<0.001

Table 5: Results from mediation analysis, for factors hypothesized to mediate the impact of SMI on financial protection indicators in Honduras; analysis is limited to financial protection outcomes which SMI was found to have significantly impacted (see Table 2)

Hypothesized causal pathways for SMI impacting the proportion of total consumption spent on prescription drugs among all households in Honduras			
Pathway	Mediating effect size	Proportion mediated	N
SMI -> usual facility is public -> proportion total expenditure on prescription drugs	-0.001** [-0.002, 0.000]	7.0%** [1.0%, 19.6%]	5019
SMI -> woman or child in HH was sick in past 2 weeks -> proportion total expenditure on prescription drugs	-0.001** [-0.002, -0.001]	10.0%** [3.6%, 23.5%)	5019
SMI -> HH member was hospitalized in past year -> proportion total expenditure on prescription drugs	-0.00 [-0.001, 0.00]	1.2% [-3.4%, 6.4%]	5019
SMI -> proportion of tracer drugs available at facility visited by HH in past month -> proportion total expenditure on prescription drugs	0.00 [-0.00, 0.00]	-0.0% [-4.8%, 3.8%]	946

Hypothesized causal pathways for SMI impacting households forgoing care due to cost in the past year in Honduras			
Pathway	Mediating effect size	Proportion mediated	N
SMI -> usual facility is public -> HH forwent care due to cost in past year	0.01** [0.001, 0.01]	-3.2%** [-7.5%, -0.52%]	5019
SMI -> woman or child in HH was sick in past 2 weeks -> HH forwent care due to cost in past year	-0.01** [-0.02, -0.004]	7.7%** [2.7%, 14.5%)	5019
SMI -> HH member was hospitalized in past year -> HH forwent care due to cost in past year	-0.001 [-0.004, 0.001]	0.5% [-1.1%, 2.3%]	5019
SMI -> proportion of tracer drugs available at facility visited by HH in past month -> HH forwent care due to cost in past year	0.00 [-0.00, 0.00]	-0.0% [-2.0%, 1.8%]	946

* p<0.05, ** p<0.01, *** p<0.001

(cnt) Table 5: Results from mediation analysis, for factors hypothesized to mediate the impact of SMI on financial protection indicators in Honduras; analysis is limited to financial protection outcomes which SMI was found to have significantly impacted (see Table 2)

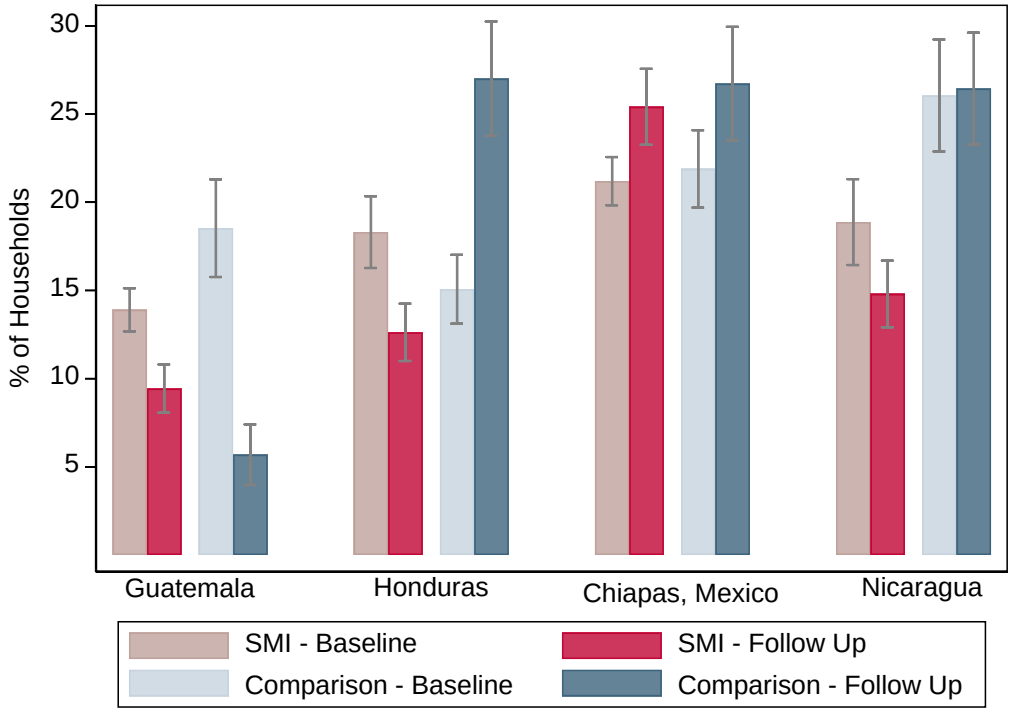
Hypothesized causal pathways for SMI impacting whether the household made any OOP health expenditures in Honduras			
Pathway	Mediating effect size	Proportion mediated	N
SMI -> usual facility is public -> HH made any OOP health expenditure	-0.01** [-0.01, -0.00]	4.3%** [0.7%, 9.3%]	5019
SMI -> woman or child in HH was sick in past 2 weeks -> HH made any OOP health expenditure	-0.01** [-0.02, -0.01]	7.2%** [3.0%, 13.0%)	5019
SMI -> HH member was hospitalized in past year -> HH made any OOP health expenditure	-0.002 [-0.01, 0.00]	1.2% [-2.2%, 4.8%]	5019
SMI -> proportion of tracer drugs available at facility visited by HH in past month -> HH made any OOP health expenditure	0.00 [-0.01, 0.01]	-0.0% [-2.0%, 1.8%]	946

Hypothesized causal pathways for SMI impacting whether the household made any OOP prescription drug expenditures in Honduras			
Pathway	Mediating effect size	Proportion mediated	N
SMI -> usual facility is public -> HH made any OOP prescription drug expenditure	-0.01** [-0.01, -0.00]	4.1%** [0.3%, 9.0%]	5019
SMI -> woman or child in HH was sick in past 2 weeks -> HH made any OOP prescription drug expenditure	-0.01** [-0.01, -0.00]	6.4%** [2.3%, 13.4%)	5019
SMI -> HH member was hospitalized in past year -> HH made any OOP prescription drug expenditure	-0.001 [-0.01, 0.00]	1.1% [-2.5%, 4.7%]	5019
SMI -> proportion of tracer drugs available at facility visited by HH in past month -> HH made any OOP prescription drug expenditure	0.00 [-0.00, 0.00]	-0.0% [-1.9%, 1.8%]	946

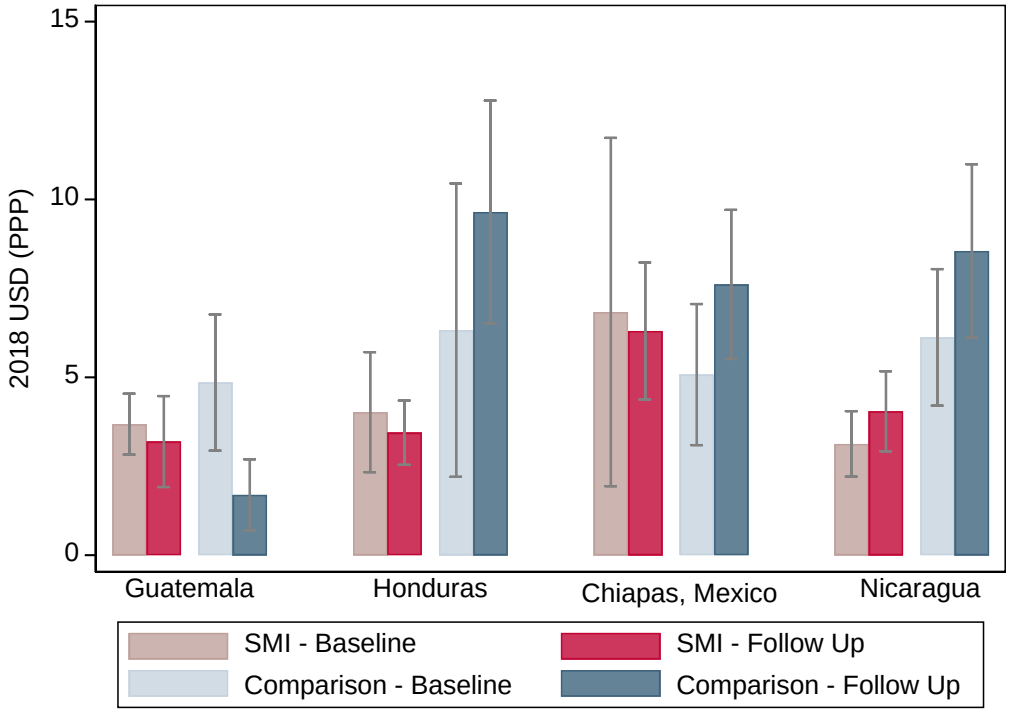
* p<0.05, ** p<0.01, *** p<0.001

Supplementary Figure 1: Secondary financial protection outcomes in SMI intervention and comparison areas at baseline and follow-up, among matched households

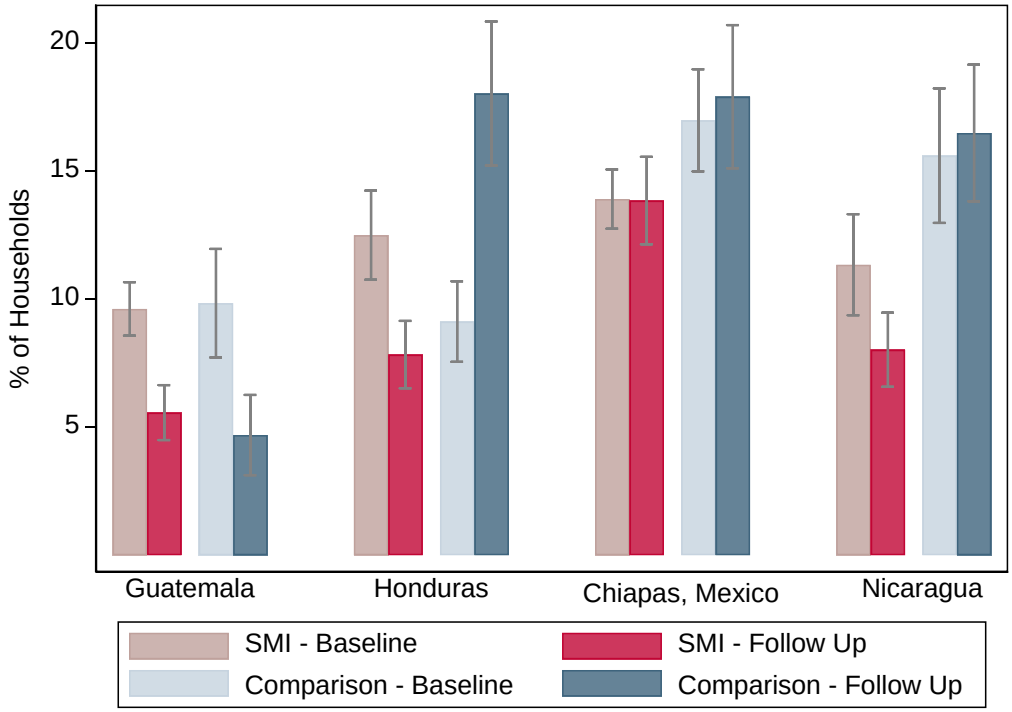
A: Percent of households making any OOP health expenditure



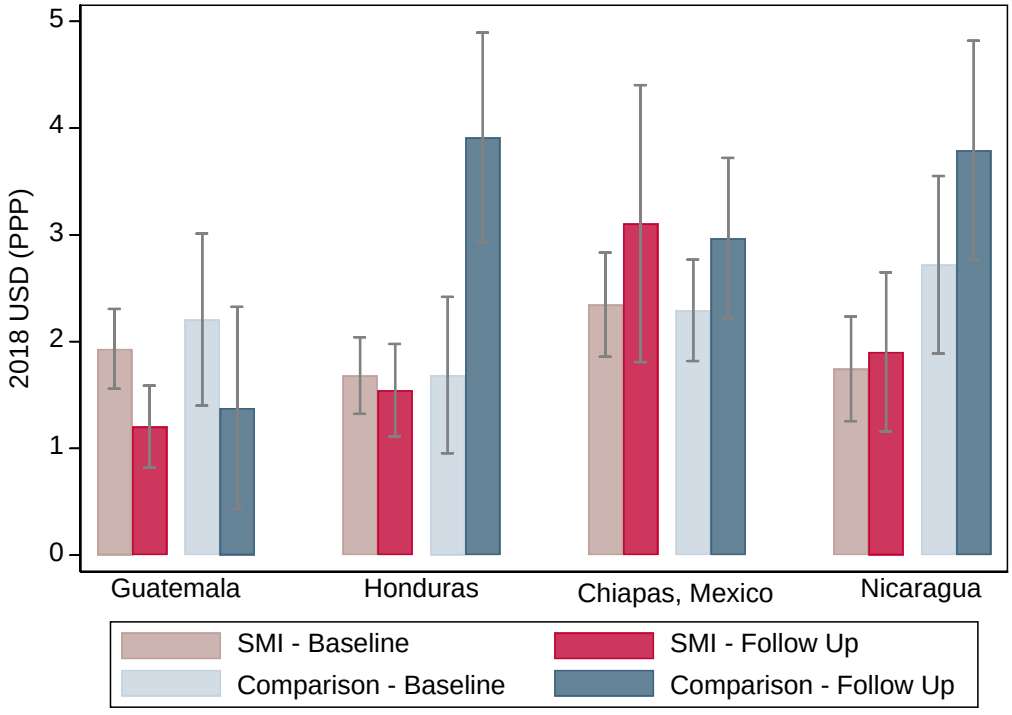
B: Mean OOP health expenditure amount (less insurance reimbursements)



C: Percent of households making any OOP prescription drug expenditure

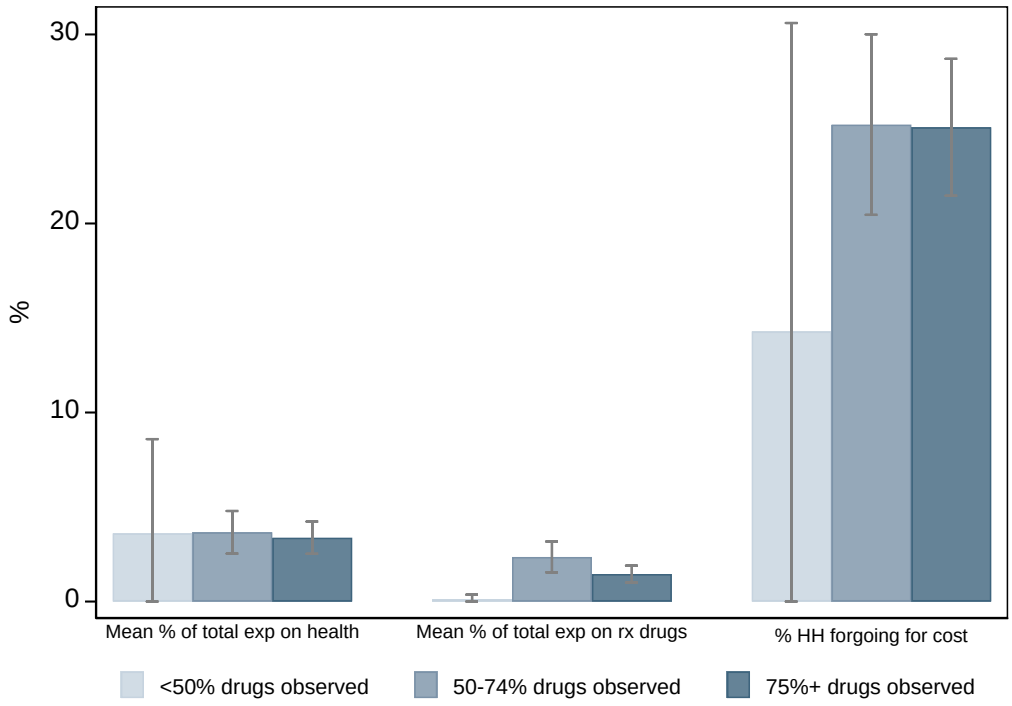


D: Mean OOP prescription drug expenditure amount

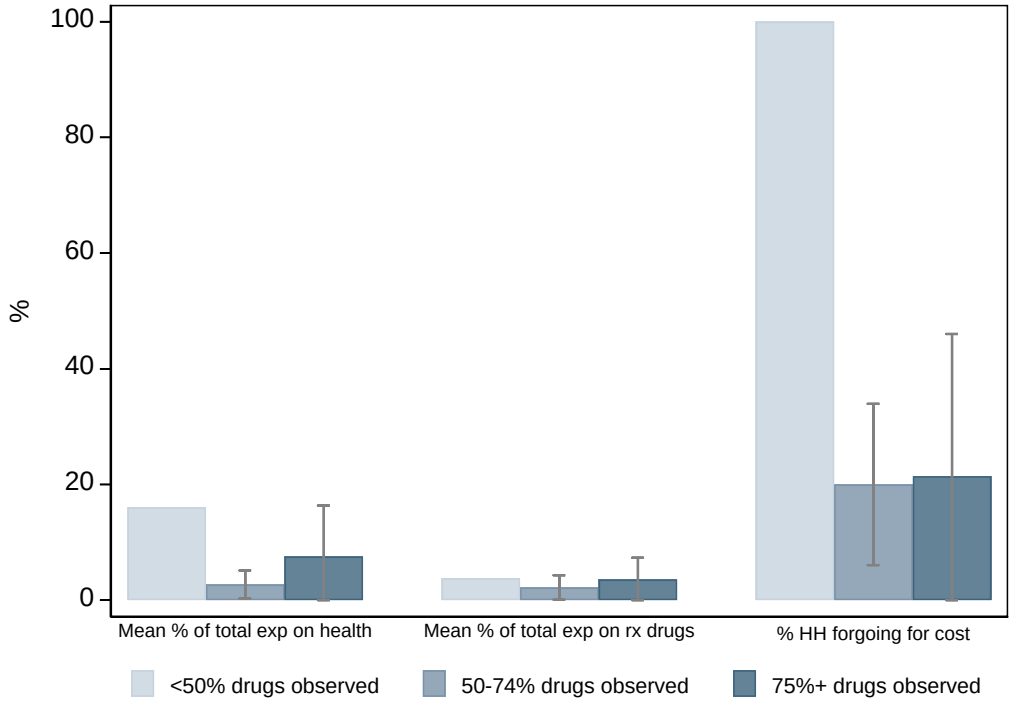


Supplementary Figure 3: Associations between drug availability and primary prescription drug expenditure outcomes in Honduras, split by ambulatory vs. basic/complete EONC-level facilities, pooled across baseline/follow-up and intervention/comparison

Ambulatory facilities

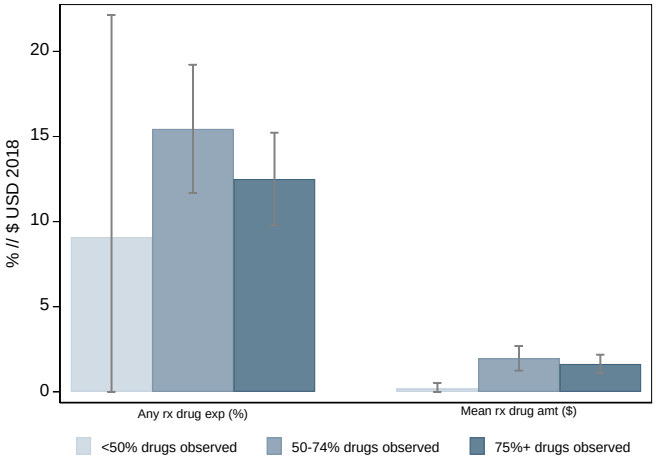


Basic/complete facilities

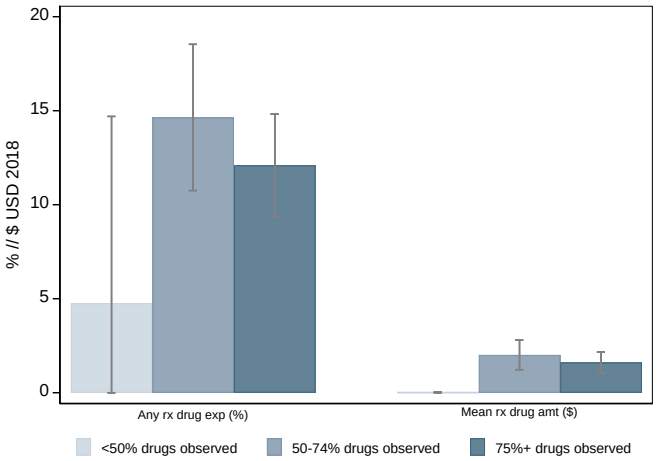


Supplementary Figure 4: Associations between drug availability and secondary prescription drug expenditure outcomes in Honduras, pooled across baseline/follow-up and intervention/comparison

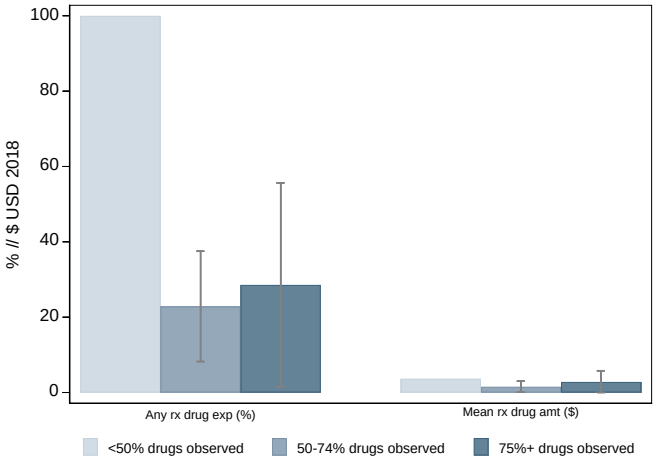
All facility types



Ambulatory facilities



Basic/complete facilities



Supplementary Table 1: List of country-specific indicator drugs, by facility EONC level																	
			Guatemala				Honduras				Mexico				Nicaragua	
Linked payment indicator	**Drug**		**Ambulatory**	**Basic**	**Complete**		**Ambulatory**	**Basic**	**Complete**		**Ambulatory**	**Basic**	**Complete**		**Ambulatory**	**Basic**
ANC/PNC	Cefalexin											X	X			
ANC/PNC	Erythromycin/Ampicillin/Penicillin										X					
ANC/PNC	Multivitamins/(Iron + Folic acid		X	X	X						X	X	X			
ANC/PNC	Nitrofurantoin										X	X	X			
Child	Albendazone / mebendazole		X	X			X	X			X	X	X		X	X
Child	Amoxicillin/Erythromycin/Benzathine penicillin						(CESAMO only)	X								
Child	Erythromycin / ampicillin / penicillin benzathine										(with doc only)					X
Child	Ferrous sulfate / micronutrients for chlidren		X	X	X						X	X	X			
Child	Ferrous sulfate /zinc sulfate / zinc gluconate														X	X
Child	Oral rehydration salts		X	X	X		X	X			X	X	X		X	X
Child	Saline / Hartmann's solution/Dextrose							X								
Child	Zinc sulfate/gluconate		X	X	X		X	X								
Child & Emergency	Antibiotics		X	X	X											X
Child & Emergency	Penicillin crystalline / ampicillin / amoxicillin											X	X			
Child (& Delivery in Chiapas)	Ringer’s lactate / Hartmann’s solution / saline solution			X	X							X	X			
Delivery	Hyoscine butylbromide/Butylscopolamine											X	X			
Delivery	Insulin syringe											X	X			
Delivery	Lidocaine without epinephrine (simple lidocaine)											X	X			
Delivery	Ophthalmic chloramphenicol drops/1% silver nitrate/oxytetracycline opthalmic											X	X			
Delivery	Povidone-iodine											X	X			
Delivery	Vitamin K/Phytonadione											X	X			
Delivery & Emergency	Ergonovine/ergometrine/oxytocin											X	X			
Emergency	Adrenaline								X							
Emergency	Amikacin sulfate												X			
Emergency	Amoxicillin/Ampicillin/Amikacin sulfate/ Crystalline penicillin G/Clindamycin/Cephalexin/Dicloxacillin /Doxycycline/Gentamicin/Metronidazole								X							
Emergency	Ampicillin							X								
Emergency	Ampicillin / crystal benzylpenicillin / penicillin / chloramphenicol				X											
Emergency	Ampicillin / penicillin			X												

(Cnt) Supplementary Table 1: List of country-specific indicator drugs, by facility EONC level																	
			Guatemala				Honduras				Mexico				Nicaragua	
Linked payment indicator	**Drug**		**Ambulatory**	**Basic**	**Complete**		**Ambulatory**	**Basic**	**Complete**		**Ambulatory**	**Basic**	**Complete**		**Ambulatory**	**Basic**
Emergency	Atropine/Atropine sulfate/Epinephrine								X							
Emergency	Calcium gluconate			X	X											
Emergency	Ceftriaxone				X								X			
Emergency	Chloramphenicol/metronidazole												X			
Emergency	Dexamethasone															X
Emergency	Dexamethasone/betamethasone			X	X				X			X	X			
Emergency	Dextrose								X							
Emergency	Diazepam				X				X							
Emergency	Diazepam/Midazolam												X			
Emergency	Diphenhydramine				X											
Emergency	Ergobasin / ergonovine maleate / ergometrine / oxytocin															X
Emergency	Ergometrine/Ergobasine maleate/Oxytocin							X	X							
Emergency	Furosemide								X				X			
Emergency	Gentamicin							X				X				X
Emergency	Gentamicin / amikacin			X	X											
Emergency	Hydralazine			X								X	X			X
Emergency	Hydralazine hydrochloride				X											
Emergency	Hydralazine/Hydralazine hydrochloride/Methyldopa/ Propanolol/Nifedipine								X							
Emergency	Isotonic crystalloids (saline solution or Ringer's lactate)								X							
Emergency	Magnesium sulfate			X	X			X	X			X	X			X
Emergency	Metronidazole / clindamycin			X	X											
Emergency	Naloxone hydrochloride								X							
Emergency	Nifedipine				X								X			
Emergency	Normal saline solution for washing								X							
Emergency	Oxytocin / methylergonovine			X	X											
Emergency	Phenobarbital/Phenobarbital sodium								X							
Emergency	Sevoflurane/propofol												X			
Emergency	Sodium bicarbonate								X							
Emergency	Succinylcholine chloride (suxamethonium)/vecuronium												X			
Emergency	Tetracycline opthalmic ointment								X							

Supplementary Table 2A: Supplementary household sample size information

Total household sample, before matching

	SMI intervention areas		Comparison areas		
	Baseline	Follow Up	Baseline	Follow Up	Total
Guatemala	3,180	1,865	796	736	6,577
Honduras	1,406	1,638	1,314	739	5,097
Mexico	3,761	1,670	1,493	746	7,670
Nicaragua	1,245	1,767	749	760	4,521
Total	9,592	6,940	4,352	2,981	23,865

Matched households that visited a health facility in the past month

	SMI intervention areas		Comparison areas		
	Baseline	Follow Up	Baseline	Follow Up	Total
Guatemala	839	493	201	187	1,720
Honduras	581	722	489	337	2,129
Mexico	1464	536	580	226	2,806
Nicaragua	592	552	421	324	1,889
Total	3,476	2,303	1,691	1,074	8,544

Supplementary Table 2B: Supplementary health facility sample size information

Ambulatory health facilities

	SMI intervention areas		Comparison areas		
	Baseline	Follow Up	Baseline	Follow Up	Total
Guatemala	47	41	21	20	129
Honduras	45	41	20	14	120
Mexico	41	46	19	22	128
Nicaragua	32	37	23	19	111
Total	165	165	83	75	488

Basic/complete health facilities

	SMI intervention areas		Comparison areas		
	Baseline	Follow Up	Baseline	Follow Up	Total
Guatemala	17	20	8	11	56
Honduras	14	13	11	11	49
Mexico	19	14	11	8	52
Nicaragua	8	23	1	11	43
Total	58	70	31	41	200

Ambulatory facilities visited at both baseline and followup

	SMI intervention areas		Comparison areas		
	Baseline	Follow Up	Baseline	Follow Up	Total
Guatemala	19	20	11	9	59
Honduras	22	18	9	8	57
Mexico	20	24	10	11	65
Nicaragua	21	18	13	11	63
Total	82	80	43	39	244

Basic/Complete facilities visited at both baseline and followup

	SMI intervention areas		Comparison areas		
	Baseline	Follow Up	Baseline	Follow Up	Total
Guatemala	17	19	7	6	49
Honduras	14	13	9	7	43
Mexico	18	14	9	8	49
Nicaragua	6	9	1	3	19
Total	55	55	26	24	160

Supplementary Table 3: Sample characteristics before and after matching

Sample charachteristics - before matching	Guatemala				Honduras				Chiapas, Mexico				Nicaragua			
	SMI Intervention areas		Comparison areas		SMI Intervention areas		Comparison areas		SMI Intervention areas		Comparison areas		SMI Intervention areas		Comparison areas	
	Baseline	Follow Up	Baseline	Follow Up	Baseline	Follow Up	Baseline	Follow Up	Baseline	Follow Up	Baseline	Follow Up	Baseline	Follow Up	Baseline	Follow Up
Household size (median)	6	6	6	5	5	5	5	5	5	5	5	4	5	4	5	5
Indigenous	79%	84%	52%	51%	0%	1%	0%	1%	79%	76%	39%	41%	18%	21%	0%	1%
Median monthly household expenditure	$ 43	$ 46	$ 43	$ 49	$ 47	$ 57	$ 64	$ 92	$ 34	$ 44	$ 47	$ 57	$ 54	$ 81	$ 61	$ 96
Max maternal education: None	30%	20%	18%	15%	7%	7%	5%	4%	15%	10%	8%	5%	9%	9%	6%	7%
Max maternal education: Primary	51%	57%	54%	58%	67%	61%	63%	52%	48%	42%	45%	32%	47%	44%	38%	32%
Max maternal education: Secondary and up	19%	23%	28%	27%	26%	32%	32%	45%	36%	49%	47%	62%	44%	48%	56%	61%
Female HoH	13%	15%	17%	14%	19%	35%	23%	40%	11%	14%	9%	17%	25%	27%	29%	30%
Any infant <1	49%	44%	54%	44%	41%	38%	43%	35%	46%	39%	47%	39%	44%	38%	42%	35%
Any elderly 60+	14%	18%	19%	18%	15%	15%	15%	15%	9%	9%	11%	11%	12%	11%	17%	17%
Has health insurance	9%	27%	19%	30%	1%	1%	2%	6%	86%	91%	86%	91%	5%	7%	11%	12%
Bono 10,000 participatoin					25%	22%	21%	13%								
Oportunidades participation									67%	63%	54%	54%				
N	3180	1865	796	736	1406	1638	1314	739	3761	1670	1493	746	1245	1767	749	760

Trimmed sample characteristics - after matching (using matching weights)	Guatemala				Honduras				Chiapas, Mexico				Nicaragua			
	SMI Intervention areas		Comparison areas		SMI Intervention areas		Comparison areas		SMI Intervention areas		Comparison areas		SMI Intervention areas		Comparison areas	
	Baseline	Follow Up	Baseline	Follow Up	Baseline	Follow Up	Baseline	Follow Up	Baseline	Follow Up	Baseline	Follow Up	Baseline	Follow Up	Baseline	Follow Up
Household size (median)	5	5	5	5	4	4	4	4	4	4	4	4	4	4	4	4
Indigenous	56%	56%	56%	56%	0%	0%	0%	0%	53%	53%	53%	53%	0%	0%	0%	0%
Median monthly household expenditure	$ 57	$ 57	$ 60	$ 60	$ 62	$ 62	$ 62	$ 62	$ 54	$ 58	$ 56	$ 53	$ 55	$ 58	$ 57	$ 56
Max maternal education: None	25%	25%	25%	25%	23%	23%	23%	23%	17%	17%	17%	17%	23%	23%	23%	23%
Max maternal education: Primary	44%	44%	44%	44%	33%	33%	33%	33%	39%	39%	39%	39%	29%	29%	29%	29%
Max maternal education: Secondary and up	31%	31%	31%	31%	43%	43%	43%	43%	44%	44%	44%	44%	48%	48%	48%	48%
Female HoH	15%	17%	20%	15%	20%	36%	28%	35%	15%	18%	9%	18%	21%	24%	30%	25%
Any infant <1	52%	39%	52%	39%	37%	33%	37%	34%	41%	38%	45%	43%	44%	39%	42%	40%
Any elderly 60+	11%	17%	16%	16%	11%	15%	16%	13%	7%	13%	9%	11%	7%	12%	15%	17%
Has health insurance	35%	35%	35%	35%	10%	10%	10%	10%	65%	65%	65%	65%	16%	16%	16%	16%
Bono 10,000 participatoin					17%	19%	15%	11%								
Oportunidades participation									46%	39%	44%	39%				
N	3079	1760	763	703	1393	1609	1295	722	3444	1576	1367	727	996	1358	737	746

Supplementary table 4: Estimated impact of SMI on household out of pocket prescription drug expenditure indicators in difference-in-differences analyses, following matching and controlling for covariates, among households that visited a health facility in the past month

	Guatemala	Honduras	Chiapas, Mexico	Nicaragua
Primary outcome				
Proportion of total household consumption spent on health	-0.05 [-0.12, 0.01]	-0.01 [-0.06, 0.03]	-0.02 [-0.06, 0.02]	-0.04 [-0.10, 0.02]
Proportion of total household consumption spent on prescription drugs	-0.05* [-0.09, -0.00]	-0.02* [-0.04, -0.00]	0.01 [-0.01, 0.03]	-0.02 [-0.03, 0.00]
Household forwent care due to cost in the past year	0.06 [-0.07, 0.19]	-0.10 [-0.23, 0.03]	-0.03 [-0.15, 0.10]	-0.00 [-0.11, 0.10]
Secondary outcomes				
Household made any OOP expenditure on health	-0.01 [-0.15, 0.12]	-0.11 [-0.24, 0.02]	-0.06 [-0.18, 0.05]	-0.11 [-0.25, 0.03]
OOP expenditure on health (2018 USD PPP)	-3.70 [-8.57, 1.18]	5.53 [-8.78, 19.84]	-2.73 [-7.38, 1.92]	-9.63 [-20.52, 1.25]
Household made any OOP expenditure on prescription drugs	-0.11 [-0.23, 0.02]	-0.11* [-0.22, -0.01]	-0.03 [-0.12, 0.07]	-0.13* [-0.24, -0.02]
OOP expenditure on prescription drugs (2018 USD PPP)	-3.73 [-7.54, 0.08]	-1.14 [-4.12, 1.84]	0.05 [-2.23, 2.33]	-4.58* [-8.96, -0.21]
N	1720	2129	2806	1889

* p<0.05, ** p<0.01, ***p<0.001

Supplementary table 5: Estimated impact of SMI on household out of pocket non-prescription drug expenditure indicators in difference-in-differences analyses, following matching and controlling for covariates

	Guatemala	Honduras	Chiapas, Mexico	Nicaragua
Among all households				
Proportion of total household consumption spent on non-prescription drugs	0.00 [-0.00, 0.01]	-0.00 [-0.01, 0.00]	0.00 [-0.01, 0.01]	-0.01 [-0.02, 0.01]
Household made any OOP expenditure on non-prescription drugs	0.04* [0.01, 0.08]	-0.05 [-0.10, 0.01]	0.00 [-0.04, 0.05]	-0.05 [-0.12, 0.03]
OOP expenditure on non-prescription drugs (2018 USD PPP)	0.24* [0.03, 0.46]	-0.02 [0.79, 0.76]	-0.57 [-1.32, 0.17]	-0.63 [-1.57, 0.32]
Among households visiting a health facility in the past month				
Proportion of total household consumption spent on non-prescription drugs	-0.00 [-0.01, 0.01]	0.00 [-0.00, 0.01]	-0.02* [-0.03, -0.00]	-0.00 [-0.01, 0.00]
Household made any OOP expenditure on non-prescription drugs	-0.00 [-0.10, 0.11]	-0.05 [-0.12, 0.03]	-0.03 [-0.11, 0.05]	-0.04 [-0.15, 0.07]
OOP expenditure on non-prescription drugs (2018 USD PPP)	0.08 [-0.51, 0.66]	0.66 [-0.81, 2.13]	-2.21 [-3.98, -0.44]	-0.46 [-1.56, 0.65]

* p<0.05, ** p<0.01, ***p<0.001

Supplementary Table 6A: Tracer drug availability at ambulatory facilities

SMI facilities	Baseline				Follow-Up			
	Guatemala	Honduras	Chiapas, Mexico	Nicaragua	Guatemala	Honduras	Chiapas, Mexico	Nicaragua
Antibiotic	0%	100%	81%	(missing)	100%	100%	100%	100%
Ferrous Sulfate	74%	71%	57%	97%	49%	95%	85%	86%
ORS	81%	93%	86%	91%	80%	100%	100%	100%
Albendazole/ Mebendazole	57%	100%	84%	94%	98%	100%	98%	100%
Folic acid	91%	93%	64%	75%	95%	95%	98%	95%
Zinc	60%	4%	22%	72%	85%	100%	30%	92%
Iron	89%	91%	27%	91%	34%	88%	89%	97%

Comparison facilities	Baseline				Follow-Up			
	Guatemala	Honduras	Chiapas, Mexico	Nicaragua	Guatemala	Honduras	Chiapas, Mexico	Nicaragua
Antibiotic	0%	17%	92%	(missing)	100%	100%	94%	100%
Ferrous Sulfate	90%	100%	32%	100%	90%	93%	41%	63%
ORS	85%	90%	79%	91%	80%	100%	100%	100%
Albendazole/ Mebendazole	80%	100%	84%	100%	100%	100%	95%	95%
Folic acid	95%	85%	57%	83%	95%	93%	91%	84%
Zinc	70%	5%	21%	91%	85%	86%	32%	79%
Iron	100%	90%	29%	87%	25%	86%	64%	95%

Supplementary Table 6B: Tracer drug availability at basic/complete facilities

SMI facilities	Baseline				Follow-Up			
	Guatemala	Honduras	Chiapas, Mexico	Nicaragua	Guatemala	Honduras	Chiapas, Mexico	Nicaragua
Antibiotic	100%	100%	65%	100%	100%	100%	100%	100%
Ferrous sulfate	76%	79%	56%	100%	100%	100%	93%	100%
ORS	100%	100%	56%	86%	100%	100%	100%	100%
Albendazole/ Mebendazole	82%	71%	61%	100%	95%	92%	93%	100%
Dexamethasone	82%	63%	13%	100%	100%	92%	93%	100%
Folic acid	94%	86%	50%	71%	95%	100%	100%	96%
Zinc	88%	7%	13%	86%	100%	100%	36%	96%
Iron	94%	86%	25%	86%	40%	100%	93%	100%
Magnesium sulfate	88%	100%	47%	86%	100%	77%	100%	96%
Oxytocin	88%	93%	53%	86%	100%	100%	100%	100%
Hydralazine	23%	75%	33%	86%	60%	100%	64%	100%
Genamycin	64%	89%	11%	80%	95%	100%	86%	100%
Diazepam	100%	100%	17%	100%	22%	100%	57%	100%

Comparison facilities	Baseline				Follow-Up			
	Guatemala	Honduras	Chiapas, Mexico	Nicaragua	Guatemala	Honduras	Chiapas, Mexico	Nicaragua
Antibiotic	100%	100%	100%	100%	100%	100%	100%	100%
Ferrous sulfate	88%	82%	40%	100%	100%	100%	88%	100%
ORS	88%	73%	100%	100%	91%	100%	100%	100%
Albendazole/ Mebendazole	63%	64%	90%	100%	73%	100%	88%	91%
Dexamethasone	63%	43%	30%	100%	91%	73%	25%	91%
Folic acid	100%	73%	100%	100%	100%	91%	100%	82%
Zinc	100%	9%	0%	100%	100%	82%	25%	82%
Iron	100%	64%	57%	0%	45%	91%	88%	82%
Magnesium sulfate	50%	82%	30%	100%	91%	64%	50%	91%
Oxytocin	88%	89%	70%	100%	100%	100%	38%	91%
Hydralazine	29%	57%	0%	100%	55%	100%	13%	91%
Genamycin	57%	86%	14%	100%	100%	80%	25%	82%
Diazepam	100%	100%	67%	(missing)	30%	100%	25%	67%

Supplementary Table 7A: Impact of SMI on alternate measures of drug availability at facilities

	Guatemala	Honduras	Chiapas, Mexico	Nicaragua
All tracer drugs (common across countries) in stock day of survey	0.12 [-0.10, 0.34]	0.04 [-0.18, 0.25]	0.15 [-0.06, 0.35]	0.34* [0.00, 0.67]
Proportion of country-specific indicator drugs in stock day of the survey	0.09 [-0.02, 0.20]	0.01 [-0.06, 0.09]	0.16** [0.05, 0.28]	0.20** [0.09, 0.32]
All country-specific indicator drugs in stock day of survey	0.12 [-0.15, 0.40]	0.28*** [0.12, 0.44]	0.58*** [0.34, 0.81]	0.37** [0.12, 0.61]
All country-specific indicator drugs in stock day of survey and 3 months prior, without interruption	0.13 [-0.10, 0.37]	0.40** [0.14, 0.67}	0.34*** [0.14, 0.54]	0.38* [0.07, 0.68]
N	182	169	180	154

Supplementary Table 7B: Association between country-specific indicator drugs measures in the health facility on the day of the survey, and prescription drug expenditure indicators

Association between having all country-specific indicator drugs available in the health facility on the day of the survey, and prescription drug expenditure indicators at households that visited that facility in the past month

	Guatemala	Honduras	Mexico	Nicaragua
Proportion of total household consumption spent on prescription drugs	-0.01 [-0.04, 0.01]	-0.00 [-0.02, 0.01]	0.02 [-0.04, 0.07]	-0.00 [-0.02, 0.02]
Household made any OOP expenditure on prescription drugs	-0.06 [-0.16, 0.04]	-0.03 [-0.08, 0.02]	-0.04 [-0.11, 0.03]	0.05 [-0.04, 0.15]
OOP expenditure on prescription drugs (2018 USD PPP)	-0.50 [-3.31, 2.31]	0.29 [-0.98, 1.56]	0.56 [-0.90, 2.03]	-2.39 [-8.64, 3.86]
N	533	946	413	388

Supplementary Table 7C: Association between having all country-specific indicator drugs continuosly available in the health facility for the three months prior to survey, and prescription drug expenditure indicators

	Guatemala	Honduras	Mexico	Nicaragua
Proportion of total household consumption spent on prescription drugs	0.00 [-0.03, 0.03]	-0.01* [-0.03, -0.00]	0.02 [-0.05, 0.09]	-0.00 [-0.02, 0.02]
Household made any OOP expenditure on prescription drugs	-0.01 [-0.14, 0.13]	-0.06 [-0.11, 0.00]	0.00 [-0.08, 0.09]	0.04 [-0.05, 0.14]
OOP expenditure on prescription drugs (2018 USD PPP)	1.40 [-2.60, 5.41]	-1.40 [-2.80, 0.01]	0.29 [-1.00, 1.59]	-1.39 [-6.60, 3.81]
N	533	946	413	388

* p<0.05, ** p<0.01, *** p<0.001

Aim 3: Impact of the results-based aid Salud Mesoamerica Initiative, phases one and two, on the timeliness and quality of antenatal care in the poorest communities in Honduras: A cluster randomized trial

ABSTRACT

Background: The Salud Mesoamerica Initiative (SMI) aims to address maternal and child health inequities in eight Mesoamerican countries through Results Based Aid (RBA). One specific goal is improving antenatal care (ANC) for the poorest women in Honduras. We evaluated the impact of SMI phases 1 (supply improvement phase) and 2 (service provision improvement phase) on ANC timeliness and quality in Honduras. **Methods:** 12 gestores in the poorest parts of Honduras were randomly assigned to receive SMI interventions and incentives or serve as a comparison area. At baseline and follow-up, we selected stratified random samples of ANC records and households in both intervention and comparison areas. Primary outcomes included the proportion of ANC records with an ANC visit recorded in the first trimester (ANC timeliness), and the proportion of ANC records that documented four ANC visits and a set of standard diagnostic and preventive services (ANC quality), both of which were tied to financial incentives in the initiative. Secondary outcomes included analogous ANC timeliness and quality indicators as measured through self-reported household survey data. We used difference-in-differences linear probability regression models, adjusted for potential confounders, to analyze the impact of SMI on ANC indicators, and tested several sensitivity analyses. **Findings:** We analyzed data on 531 pregnancies from ANC records (397 intervention, 134 comparison) and 2,142 pregnancies from household surveys (1,267 intervention, 875 comparison). The primary outcomes of ANC timeliness and quality as recorded in ANC records increased by 30.5 and 70.9 percentage points, respectively, in intervention areas; and 9.4 and 17.4 percentage points, respectively, in comparison areas. The adjusted difference-in-differences analysis found that SMI significantly improved both timeliness (coefficient: 0.22, 95% confidence Interval (CI): 0.01, 0.42) and quality (coefficient: 0.53, 95% CI: 0.23, 0.84) as recorded in ANC records. In sensitivity analyses, the magnitude and direction of SMI's impact in medical records was consistent for quality, but inconsistent for timing. The secondary outcomes of timeliness and quality as measured in self-reported household survey data were not significantly improved by SMI. **Interpretation:** We found early indications that the first and second phases of the SMI RBA program have improved ANC timing and quality in the marginalized populations targeted by SMI, though the possibility that results are driven by improved record-keeping practices should be investigated. Population-level outcomes should be rigorously evaluated following SMI's ongoing demand-generation phase to ensure that service improvements are benefiting all marginalized women.

INTRODUCTION

Mesoamerica, the region spanning southern Mexico to Panama, has achieved significant health gains in the past decade, but remains one of the least equitable regions in the world in terms of economic and health outcomes [1,2]. The Salud Mesoamerica Initiative (SMI) aims to address maternal and child health inequities in eight Mesoamerican countries through a Results Based Aid (RBA) model [3]. Under SMI, the Ministries of Health (MOH) from each country are eligible to receive financial incentives for improving specific indicators in their poorest populations. The initiative is funded by the Bill and Melinda Gates Foundation, the Carlos Slim Foundation, and the Spanish Agency for International Development Cooperation, and administered by Inter-American Development Bank (IDB).

Honduras is one of the eight countries participating in SMI. Honduras is a lower-middle-income country in Central America with a GDP of $2,361 per capita and a 73 year life expectancy [4]. It has a maternal mortality ratio of 109.6 per 100,000 live births, twice the regional average for Central Latin America of 55.9 per 100,000 [5]. Economically, Honduras is one of the most inequitable countries in the world, with a Gini index of 50.0 [4].

Despite a nominal policy of universal government-provided health care, coverage of health services is highly inequitable in Honduras, including for antenatal care (ANC), a preventive health intervention shown to benefit both mothers and infants [6–9]. In 2011-12, 79% of Honduran women in the poorest quintile received four ANC visits compared to 96% in the wealthiest quintile [10]. Therefore, improving ANC was identified as a key objective of SMI in Honduras. The Honduran MOH committed to meeting two targets for improving ANC timeliness and quality in SMI intervention areas, which were tied to financial incentives. In 2018, at the end of SMI's 2nd operational phase, Honduras met both of the ANC-related targets (as well as nine other targets unrelated to ANC) and therefore qualified for the RBA financial incentives.

While these indicators have improved to the target levels in SMI areas, the causal relationship between SMI and ANC services has not yet been evaluated. A number of factors other than SMI may have influenced ANC during this period, including economic growth, migration, and temporal trends of improving ANC services [10,11]. Therefore, a rigorous impact evaluation is needed to assess the causal impact of SMI's RBA design on ANC outcomes. In this study, we aim to evaluate the impact of SMI on the timeliness and quality of ANC in Honduras, following SMI's 2nd operational phases.

METHODS

Initiative and study design

SMI began with IDB and Honduras MOH conducting an analysis to identify the nine (of 18 total) departments with the greatest poverty and health inequalities. Within these nine departments, IDB randomly assigned 16 communities of similar poverty level into either treatment (eight) or control (eight). Due to budget constraints, two were dropped from each group, leaving six per group [12].

In the intervention communities, the MOH worked with the IDB and technical assistance providers to design and implement evidence-based interventions to address the highest priority maternal and child health issues. The interventions are rolled out in phases: phase 1 (2012-2014) focused on supply factors, phase 2 (2015-2017) on service provision, and phase 3 (currently underway) will target demand generation. At the end of phase 2, Honduras became eligible for results-based payments if they reached 11 pre-defined targets that were negotiated by the MOH and IDB at baseline. Two of these targets were related to ANC:

- Increase the proportion of ANC records where the first recorded visit occurred during the first trimester (before 12 weeks' gestation) from 59.2% to 63.1%
- Increase the proportion of ANC records with at least four ANC visits and a set of specific best practices recorded from 23.2% to 33.7%,

Additional incentives will be available following phase 3. The targets were set based on literature reviews of intervention effectiveness studies, trend analysis from the Global Burden of Disease study, expert advice, and a cost-benefit analysis. All indicator measurements are performed by an independent evaluator (the authors of this study).

The comparison communities were not part of target setting and did not receive any financial incentives based on performance. Some SMI interventions were targeted only to SMI areas, such as support for monitoring and continuous quality improvement. However, other interventions initiated under SMI could not feasibly be targeted to only the comparison areas, including changes to country norms.

Sample

This evaluation used data from two sources: (a) ANC records from ambulatory facilities (Centro de Salud Rural (CESAR) and Centro de Salud con Medico y Odontologo (CESAMO) - the primary provider of ANC in Honduras) and (b) household surveys of women. Stratified random samples of ANC records and women from households were selected at baseline in 2013, and at follow-up in 2017 (following the end of SMI's 2nd phase). Baseline and follow-up samples were selected separately and sample sizes were powered to detect changes in the payment indicators [12].

ANC record sample: The first stage of ANC record sampling was to select ambulatory health facilities from an MOH-provided frame of all ambulatory facilities in the SMI and comparison areas. At baseline, we randomly selected ambulatory facilities in SMI and comparison communities. At follow-up, the sampling was more complex, as we selected half the facilities from the sample of facilities selected at baseline, and the other half from facilities not selected at baseline. Within each of these stratum (baseline repeat and non-baseline repeat), we first sampled from among facilities that were in the same gestor as households selected in the follow-up sample (see below for household selection process). If there were not enough ambulatory facilities from communities in the household sample to meet the quota, we then randomly sampled from among ambulatory facilities in the remaining communities.

Within the selected facilities, there was a second stage of randomly selecting ANC records. Inclusion criteria for the ANC records were that it was an un-complicated pregnancy that resulted in birth (as ANC norms would differ for a complicated pregnancy that required more visits, or you would not expect the

same visits and tests for a pregnancy that resulted in a miscarriage) occurring in the two years prior to the survey and entirely during the intervention period.

Household women sample: The household sampling frame included all census tracts (our Primary Sampling Unit (PSU)) located in SMI and comparison areas, based on the 2001 national census (the most recent available). We randomly sampled PSUs within the SMI and comparison strata, using probability proportional to the number of households. For each selected PSU, we conducted a new census of households and residents to account for population movement since the last government census. We then identified eligible households (those with women 15-49 years or children <5) based on our census and randomly selected 30 eligible households in each PSU to be part of the survey. If multiple women were present in a household, the survey was administered to the woman who most recently had her birthday. Separate household samples were selected at baseline and follow-up to avoid introducing observer bias by re-visiting the same households or communities; the same procedure was used to select both samples.

For this analysis, inclusion criteria for the household survey included women 15-49 years old with a pregnancy in the two years prior to data collection. To minimize recall bias, we only included the most recent pregnancy in this two year period.

Data collection

All data were collected by trained personnel using computer-assisted personal interview (CAPI) with the DatStat Illume software. Data were uploaded daily to a central server, allowing continuous data quality monitoring and feedback. The study received institutional review board (IRB) approval from the University of Washington, partnering data collection agencies, and the Ministry of Health. Informed consent was obtained from all women and health facility directors prior to data collection.

From the ANC records, medical professionals trained as data collectors for our study extracted information about each ANC visit, including the gestational week, and whether specific diagnostics and test were performed at the visit, for pregnancies during the past two years.

The household survey consisted of several modules. The first collected information about the household, including wealth and characteristics of the home. The second was a maternal health questionnaire containing demographic, health behavior and reproductive health information. Women reported whether they had received any ANC, the total number of visits, gestational week or month of their first ANC visit, and whether they had received certain tests and diagnostics at any point during their most recent pregnancy during the past two years.

Primary outcomes

Our primary outcomes were the indicators tied to financial incentives at the end of SMI's second phase, both of which were measured through ANC records:

- *Timeliness, measured in ANC records:* Proportion of ANC records where the first recorded visit occurred during the first trimester (before 12 weeks' gestation)

- *Quality, measured in ANC records:* Proportion of ANC records with at least four ANC visits and the following best practices recorded:
    - Best practice physical checkups performed at each visit: weight, blood pressure, fundal height (if gestational age >=22 weeks), fetal heart rate (if gestational age >20 weeks), fetal movement (if gestational age >20 weeks).
    - Best practice lab tests performed at least once: blood glucose level, HIV test, blood type (Rh factor, blood group), urinalysis (proteinuria), anemia (Hb level), and syphilis (VDRL/RPR).

We also tested using an alternate continuous quality specification, measured as the proportion of tests/procedures performed rather than as a binary outcome.

Secondary outcomes

Our secondary outcomes were analogous to the primary outcomes, but measured by women's self-report in the household surveys. These indicators were not tied to financial incentives, but indicators similar to these[1] were identified by SMI as important non-financial monitoring indicators.

- *Timeliness, self-reported:* Proportion of women who self-reported attending their first ANC visit during the first trimester (before 12 weeks' gestation).
- *Quality, self-reported*: Proportion of women who self-reported receiving at least four ANC visits, and best practice physical check-ups and lab tests at any time during the pregnancy. This included the same best practices and lab tests that comprised the ANC records indicator, except fetal heartbeat which was erroneously left out of the household survey. Due to limitations of recall, the household indicator measured whether these practices and tests had occurred at any time during the pregnancy, as opposed to the more specific timing and frequency requirements applied to the ANC record indicator.

We included an additional secondary outcome to measure whether women's utilization of ANC at public facilities had been increased by SMI's supply and service delivery interventions.

- *Utilization, self-reported:* Proportion of women who self-reported attending any ANC at a public health facility

Statistical Analyses

Descriptive analysis: We conducted descriptive analyses to compare the baseline and follow-up values for each indicator, analyze changes in the distribution of the timing of visits, and examine the sub-components of the quality indicators.

Difference-in-differences analysis: We used a difference-in-differences (DID) framework to evaluate the impact of SMI on the primary and secondary outcomes. The first difference was between intervention and comparison areas and the second was between the baseline and follow-up time points. We

[1] The official SMI ANC household monitoring indicators incorporate an additional criterion of receiving the ANC from a 'skilled' provider. We ignored this criterion for the purpose of this manuscript to allow comparability with the primary outcomes from the ANC records, which had no requirements about provider type. This minimally affected the results, as 92% of women receiving ANC reported seeing a 'skilled' provider.

estimated the impact of SMI using linear probability regression models with the following basic form [13]:

$$ANC\ outcome = \beta_1 Follow\text{-}up + \beta_2 SMI + \beta_3 Follow\text{-}up*SMI + \beta_4 X + \varepsilon$$

Where *Follow-up* is a binary indicator of the time period (baseline=0, follow up=1); SMI is a binary indicator of whether the woman was from an intervention or comparison area (comparison=0, intervention=1); Follow-up*SMI is an interaction of these two binary indicators, taking on a value of 1 for women in SMI intervention areas in the follow-up period and 0 otherwise; and X is a vector of covariates. The coefficient of interest that measures the impact of SMI is β_3.

The main model for the primary outcomes measured through ANC records used the record as the unit of analysis; accounted for the clustering of records within health facilities; and controlled for women's education level (the only potential confounder available from the record extraction).

The main model for the secondary outcomes measured through household surveys used the woman as the unit of analysis; accounted for sample weights and the clustering of women within survey segments; and controlled for factors that were significantly associated with ANC timeliness and/or coverage in Honduras in a previously published analysis of SMI baseline data [14]. These potential confounders included: household wealth quintile as compared to other SMI households, number of children, education level, mobile phone ownership, media exposure in the past month, and previous obstetric complications. We also controlled for participation in the Bono 10,000 conditional cash transfer program, which provides incentives for pregnant women to attend medical care.

This was an intention to treat analysis, with women classified based on whether their gestor had been assigned to the SMI intervention or comparison group. Per-protocol or completer analyses were not relevant given the macro nature of the RBA intervention, and because detailed information was not collected about how and where specific interventions were implemented.

Sensitivity analyses: We conducted several sensitivity analyses (all within the DID framework) to check if the results from our main model were consistent with alternate specifications (Supplemental Box 1).

For the ANC records, these included a two-sided two-sample t-test with the record as the unit of analysis; a two-sided two-sample t-test where outcomes were collapsed to the facility level before being compared; a regression analysis like our main specification that included clustering but without controlling for confounders; a regression analysis like our main specification but restricted to 208 records coming from the 27 facilities sampled at both baseline and follow-up; a two-sided two-sample t-test of outcomes collapsed to the facility level, only among the 27 facilities sampled at both baseline and follow-up; and an alternate specification of the quality indicator measuring the proportion of tests/procedures received rather than a binary indicator of if all standards were met.

For the household survey data, sensitivity analyses included a two-sided two-sample t-test with the woman as the unit of analysis; a two-sided two-sample t-test where outcomes were collapsed to the gestor level before being compared; a regression analysis like our main specification that included clustering but without controlling for confounders; and a matched women analysis, where the sample

was restricted to only include women with a socio-demographic profile that matched to at least one woman in each of the other samples (intervention-baseline, intervention-follow-up, comparison-baseline, comparison-follow-up). More details on this matched women analysis are provided in Supplemental Box 1 (note that unlike Aim 2, which included three countries that did not have random assignment of communities, we did not feel the matched sample needed to be the primary model specification for Aim 3, given that Honduras had random assignment of communities). We also tested an alternate specification of the quality indicator measuring the proportion of tests/procedures received rather than a binary indicator of if all standards were met.

Sub-group analyses: We estimated the treatment effects within several sub-groups including women with primary/pre-primary education; women with secondary or more education; women in the poorest quintile relative to other sampled women; and women in the wealthiest quintile relative to other sampled women. The wealth sub-group analyses were only possible for the household-based results, as this information was not contained in the ANC records.

Additionally, we estimated the treatment effects for the household-based timeliness and quality outcomes among the 82% of women in our household sample who received ANC at a CESAR or CESAMO facility. This was done to align the women making up the household data with the ANC records data (which also came from CESAR and CESAMO facilities), allowing for a more direct comparison of the record-based and household-based results, as they would theoretically be measuring the same population of women.

All analyses were conducted using Stata version 13.1.

RESULTS

Sample

A total of 550 ANC records meeting inclusion criteria were reviewed and extracted; after removing 19 records missing information on one or more variable necessary for analysis, our sample for analysis included 531 ANC records (125 baseline intervention, 272 follow-up intervention, 58 baseline comparison, 76 follow-up comparison) from 93 facilities (Table 1). The median age was 23 years, and most (89%) women were married. In total, 4% had no education, 85% primary, 11% secondary, and <1% university; this was similar across the time periods, and intervention and control groups. Characteristics were similar when looking only among the 27 repeat facilities (Supplemental Table 1).

A total of 2,184 women meeting inclusion criteria were interviewed in household surveys; after removing 42 women missing information on one or more variable necessary for analysis, our sample for analysis consisted of 2,142 women (589 baseline intervention, 678 follow-up intervention, 573 baseline comparison, 302 follow-up comparison) (Table 1). Compared to the ANC records, the women in the household sample were slightly older (median 25 years), and less likely to be married (79%) and literate (72%). Median household monthly expenditure was higher in the comparison group at both baseline (121 USD among intervention, 161 USD among comparison) and follow-up (115 USD among

intervention, 197 USD among comparison). In the intervention areas, 10% of women had a baccalaureate or university-level education at both baseline and follow-up, compared to 15-17% in the comparison areas. From baseline to follow-up, the proportion of women reporting a female head of household increased by 15 percentage points (PP) in intervention areas and 18 PP in comparison areas. Mobile phone ownership was approximately 10 PP lower in intervention than comparison areas and did not change over time. Media exposure was similar at baseline between intervention (81%) and comparison (79%) areas, but declined in the intervention group to 67%, while rising slightly in comparison areas to 81%. The proportion of women recently interacting with a community health worker increased in intervention areas (11% to 15%) and decreased in comparison areas (13% to 6%). Prior obstetric complications (11% overall) and satisfaction with a recent facility visit (95% overall) were similar across the groups. Around 20% of women participated in the Bono 10,000 conditional cash transfer program in the baseline intervention, follow-up intervention, and baseline comparison groups, but the proportion was lower (11%) in the follow-up comparison group.

Primary outcomes

Timeliness, measured in ANC records: The proportion of ANC records where the woman received ANC within the first trimester grew by 30.5 PP in the intervention areas (59.2% to 89.7%, surpassing the incentive target of 63.1%), compared to a 9.4 PP increase in comparison areas (60.3% to 69.7%) (Table 2, panel facilities in Supplemental Table 2). Further, in intervention areas, the median gestational age of the first visit recorded in ANC records declined from 2.6 to 1.8 months; the change was minimal in comparison areas (2.5 to 2.3 months) (Figure 1a).

The main DID analysis found that SMI had a large and significant effect on the proportion of records with a visit recorded in the first trimester (Coefficient: 0.22, 95% Confidence Interval (CI): 0.01 to 0.42) (Table 3). This finding was consistent across 3 of the 5 sensitivity analyses, but inconsistent in the models restricted to the 208 records from 27 facilities sampled at both baseline and follow-up. Specifically, in these two models restricted to repeat facilities, the coefficients were small and the effects insignificant, for both the record-level model (Coefficient: 0.02, 95% CI: -0.22, 0.26) and when indicators were collapsed and analyzed at the facility-level (Coefficient: -0.05, 95% CI: -0.39, 0.28). We noted these repeated facilities were more often hospitals rather than outpatient clinics, and that these facilities had higher baseline values, perhaps explaining why no effect was observed. Magnitude and direction of effects were similar when stratified by educational status.

Quality, measured in ANC records: The ANC records quality indicator more than tripled in the intervention group (23.2% to 94.1%, surpassing the incentive target of 33.7%) and also increased in the comparison group (43.1% to 60.5%) (Table 2). In the intervention areas, every test was received by 99% of women at follow up. In comparison areas, most tests were received by over 90% of women, with the exception of the Hb test (68% of women) (Figure 2a).

The main DID model found that SMI had a large and statistically significant positive effect on the quality indicator (Coefficient: 0.53, 95% Confidence Interval (CI): 0.23 to 0.84) (Table 3). The coefficient was of similar direction and magnitude in most sensitivity analyses (range from 0.33 to 0.53), though the coefficients lost statistical significance when restricted to records from facilities measured at both

baseline and follow-up (likely due to reduced power with the smaller sample size). Impact was significant also when using a continuous measure of quality (% of tests/procedures performed) (Coef: 0.13, 95% CI: 0.00, 0.26) (Supplemental Table 3). Magnitude and direction of effects were similar when stratified by educational status.

Secondary outcomes

Timeliness, women's self-report: The proportion of women who reported receiving ANC within the first trimester increased from 47.2% in 63.0% (15.8 PP change) in the intervention areas, and from 55.7% to 64.6% in the comparison areas (8.9 PP change) (Table 2; matched results in Supplemental Table 2). The median month of self-reported first visit declined from 3 to 2 in intervention areas, and remained at 2 in comparison areas (Figure 1b).

We found no significant impact of SMI on self-reported ANC timeliness in the main DID model nor any of the sensitivity or sub-group analyses (Table 3, Supplemental Table 3). The coefficients for the impact of SMI on timeliness across sensitivity analyses ranged from -0.03 to 0.07, and were not statistically significantly different from zero. Results were also insignificant when restricting to women in our household sample who had received ANC at a CESAR or CESAMO facility (Supplemental Table 3).

Quality, women's self-report: The proportion of women who self-reported receiving ANC per best practices increased substantially in both intervention (24.7% to 48.1%) and comparison (22.0% to 40.1%) areas (Table 2). Reported coverage increased for all individual quality components in both intervention and comparison areas (Figure 2b). The largest increases were for laboratory tests, many of which increased by over 10 or 20 PP in both intervention and comparison areas. At follow-up, the largest remaining gap was syphilis tests, which 66% and 61% of women reported receiving in intervention and comparison areas, respectively.

The DID analyses did not find any significant effect of SMI on self-reported ANC quality, including in the sensitivity and sub-group analyses (Table 3, Supplemental Table 3). The coefficients for the impact of SMI were mostly positive (ranging from -0.04 to 0.10), but not statistically significantly different from zero. Results were also insignificant when restricting to women who had received ANC at a CESAR or CESAMO facility, with the exception of the simple woman-level t-test (Coefficient: 0.11, 95% CI: 0.01, 0.20) (Supplemental Table 3).

Utilization, women's self-report: The proportion of women who self-reported receiving any ANC at a public facility increased from 87.8% to 94.1% in intervention areas and declined from 83.2% to 80.1% in comparison areas (Table 2).

SMI did not have a significant impact on self-reported utilization in the main DID model (Coefficient: 0.05, 95% CI: -0.03, 0.14], though was significant in the simple woman-level t-test (Coefficient: 0.09, 95% CI: 0.04, 0.15) (Table 3). The sub-group analysis found that SMI had a significant impact among women in the poorest wealth quintile, both in the main adjusted DID model (Coefficient: 0.26, 95% CI: 0.14, 0.25) and in sensitivity analyses, including the matched women DID specification (Coefficient: 0.11, 95% CI: 0.01, 0.21) (Supplemental Table 3).

DISCUSSION

We find early indications that SMI has significantly improved the financially-incentivized outcomes of ANC timeliness and quality recorded in ANC records in Honduras. A positive result would reinforce previous interim evidence that showed SMI making large gains in process indicators including supply availability and provider trainings [15]. However, the impact of SMI on ANC timing is inconsistent in some sensitivity analyses, and neither timeliness nor quality were significantly impacted by SMI in the population-representative household survey data.

Our study adds to a body of literature showing promising but mixed evidence for incentive-based programs improving development and health outcomes [16]. For ANC specifically, this includes an RCT in Honduras that showed household incentives could increase ANC coverage [17] and the Rwanda Community PBF Program, which demonstrated the impact of health care provider incentives on ANC quality and demand-side incentives on ANC timeliness [18,19]. However, SMI is the first program we know of to take an RBA approach to improving ANC, meaning that it provides incentives at the government, rather than the provider or household level.

Only a few RBA programs have been implemented to date, and even fewer evaluated, including a Norwegian program to reduce deforestation in Brazil, a secondary education program in Ethiopia, a several development programs sponsored by the United Kingdom's Department for International Development [20]. These experiences have demonstrated that RBA can improve outcomes, but also raised concerns about potential unintended consequences, including (a) detracting resources from non-incentivized disease areas or geographies and (b) incentivizing data manipulation, as was shown to occur under the Gavi Alliance's Immunization Services Support (ISS) program [21–23].

It appears unlikely that SMI detracted resources from non-intervention areas, as nearly every indicator increased in comparison areas, sometimes quite significantly. In terms of the effect on other health areas, SMI incentivized a wide array of maternal and child health indicators, all of which increased to target levels. However, we do not have the data to assess the impact on non-maternal and child health areas.

We do not suspect that the results are driven by data manipulation. However, this possibility should be carefully considered, given our finding that SMI improved outcomes in the ANC records but not the household data (including among the 82% of women in the household survey who had received ANC at the same types of public facilities comprising our facility sample). Learning from the ISS experience, SMI aimed to minimize the potential for data falsification by measuring indicators through an independent evaluator, but falsification could still be introduced if providers were incentivized to record false data [12]. While theoretically plausible, we believe this is unlikely based on examining the data. First, the financial incentives were provided to the government, as opposed to the providers recording the information. We do not detect unusual pattern or distributions in the timing of the first visit that would suggest manipulation [24]. Additionally, we observed large discrepancies between the proportion of households and ANC records reporting receiving specific ANC services and tests at baseline, so the disconnect appears to pre-date SMI. Therefore, we believe that more nuanced measurement issues, not

intentional manipulation, are driving the discrepancies between record- and household-based indicators.

First, spill-over may have occurred between the intervention and control areas. Data from the baseline SMI study show that Honduran women are willing to bypass their nearest facility in order to attend a higher quality facility further away for delivery services [26]. If this is occurring for ANC services, it could mean that women from comparison areas are traveling to access higher-quality services in the SMI areas, therefore improving population-level indicators in both intervention and comparison areas, and tending towards a null result in the household data. Furthermore, some interventions, such as changes to norms could not be targeted exclusively to SMI areas, additionally biasing towards a null finding.

Another factor is that the household survey data are self-reported, introducing potential biases and greater measurement error as compared to the medical records. Women tend to heap their self-reported first visit date by months, compared to ANC records documenting precise gestational weeks (Supplemental Figure 1), and are only asked if they received services and tests at any time during the pregnancy (compared to the ANC records documenting detailed information about the timing and frequency of each service/test). As a result, the household indicators are less precise than the record-based indicators, and therefore less sensitive to change. Additionally, even in the baseline data, there are important discrepancies between the self-reported and record-based data that reduce our confidence in the validity of the household measures. The direction of bias is inconsistent across different elements of the quality indicator – some tend to be biased up in the household data as compared to documented evidence from medical records (blood test, weight, blood pressure) while others are biased down (blood glucose, HIV, syphilis). This supports our baseline assumption that medical records are a better source of information about specific tests: a woman may recall having her blood drawn, but not know if it was to test for HIV, anemia, syphilis, or something else. Additionally, self-report is subject to social desirability bias, and women in Honduras might be especially likely to over-report receiving ANC due to the incentives and culture introduced by the Bono 10,000 program [25]. Systematic over-reporting in both intervention and comparison groups and random measurement error would both lead to smaller effect sizes and lower confidence when measuring SMI's impact. For these reasons, the lack of a significant result in the household data does not necessarily reduce the credibility of the medical record-based findings.

However, the primary challenge with the ANC records is that it is difficult to differentiate between improvements in ANC service provision and improvements in record keeping practices [27,28]. SMI's focus on measurement may have led providers to be more diligent in recording each visit, and the specific procedures and tests that were performed, without improving the care itself. For instance, household data show that most women received a blood test at baseline; perhaps healthcare workers were simply not recording each of the specific tests done with blood samples at baseline, but recorded them more diligently at follow-up. This issue should be investigated through qualitative interviews with ANC providers before drawing definitive conclusions about the impact of SMI on these outcomes.

These measurement challenges in the household and medical record data are the primary limitation of our study. We also did not account for multiple comparisons, meaning we may be over-estimating significance; conversely, the study was not adequately powered for some of the sensitivity analyses,

which would under-estimate significance. Finally, this study was conducted in a setting and time where only four ANC visits were recommended; the WHO now recommends eight visits, and we do not know if the results will remain the same under the new standards.

Our study identifies several next steps for implementing and evaluating the SMI RBA model. First, qualitative work is needed to investigate if SMI has truly improved ANC timeliness and quality in CESAR and CESAMO facilities, or simply improved record keeping practices.

Second, we must monitor how SMI's ongoing demand-generation phase influences ANC utilization rates. It is often hypothesized that improving quality of services will drive demand, and we did find that the supply and service delivery phases of SMI increased utilization, though only among the very poorest households in these already marginalized areas. It is important that coverage increases alongside service quality, else we risk perpetuating inequalities between those accessing the health system (for whom services improve) and those not (who continue to be left out).

Finally, the results from this macro evaluation of the overall SMI program in Honduras should be combined with (a) impact evaluations of specific programmatic interventions deployed in Honduras in order to identify which had an effect, and (b) implementation science work to unpack which elements of SMI's design and execution drive impact. An initial process evaluation conducted in Chiapas found that SMI's success depends on a number of factors, including the financial incentives, financial resources, accountability introduced by external measurement, technical assistance, and "healthy competition" among countries participating in the initiative [29]). It remains to be seen if the conditions underlying SMI could be replicated elsewhere, and if our findings are generalizable.

Addressing these questions is necessary to verify the effectiveness of SMI and determine if and how the SMI RBA approach can be adapted to other settings and health services.

CONCLUSION

We find early indications that the first and second phases of the SMI RBA program have improved ANC timing and quality in the marginalized populations targeted by SMI, though the possibility that results are driven by improved record-keeping practices should be investigated. Population-level outcomes should be rigorously evaluated following SMI's ongoing demand-generation phase to ensure that service improvements are benefiting all marginalized women.

ACKNOWLEDGEMENTS

We thank the women and health care workers who gave their time and assistance to make this study possible. We also thank the many individuals involved in collecting the data at baseline (Fundación FES) and follow-up (UNIMER).

FIGURES

TABLES

CITATIONS

1. Belizán JM, Cafferata ML, Belizán M, Althabe F. Health inequality in Latin America. Lancet Lond Engl. 2007;370:1599–600.

2. Barreto SM, Miranda JJ, Figueroa JP, Schmidt MI, Munoz S, Kuri-Morales PP, et al. Epidemiology in Latin America and the Caribbean: current situation and challenges. Int J Epidemiol. 2012;41:557–71.

3. Regalia F, Eichler R, Gigli S, Cleason M, Tapia Conyer R, Hernandez G, et al. Not Your Cookie-Cutter Results-Based Aid Initiative: Salud Mesoamerica Initiative's Experience Improving Health for the Poorest in Mesoamerica [Internet]. Inter-American Development Bank; 2017 Oct. Available from: http://publications.iadb.org/handle/11319/8567

4. World Bank. Honduras Data [Internet]. [cited 2018 Jul 16]. Available from: https://data.worldbank.org/country/honduras

5. Kassebaum NJ, Barber RM, Bhutta ZA, Dandona L, Gething PW, Hay SI, et al. Global, regional, and national levels of maternal mortality, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. The Lancet. 2016;388:1775–812.

6. Bhutta ZA, Das JK, Bahl R, Lawn JE, Salam RA, Paul VK, et al. Can available interventions end preventable deaths in mothers, newborn babies, and stillbirths, and at what cost? Lancet. 2014;384:347–70.

7. Carroli G, Rooney C, Villar J. How effective is antenatal care in preventing maternal mortality and serious morbidity? An overview of the evidence. Paediatr Perinat Epidemiol. 2001;15 Suppl 1:1–42.

8. Abdal Qader MA, Badilla I, Mohd Amin R, Ghazi HF. Influence of antenatal care on birth weight: a cross sectional study in Baghdad City, Iraq. BMC Public Health. 2012;12:A38.

9. Beeckman K, Louckx F, Downe S, Putman K. The relationship between antenatal care and preterm birth: the importance of content of care. Eur J Public Health. 2013;23:366–71.

10. Republica de Honduras. Encuesta Nacional de Demografia y Salud: ENDESA 2011-12. 2013.

11. SS/Honduras S de S-, INE/Honduras IN de E-, International M. Honduras Encuesta Nacional de Demografía y Salud 2005-2006. 2006 [cited 2018 Jul 17]; Available from: http://dhsprogram.com/publications/publication-fr189-dhs-final-reports.cfm

12. Mokdad AH, Colson KE, Zúñiga-Brenes P, Ríos-Zertuche D, Palmisano EB, Alfaro-Porras E, et al. Salud Mesoamérica 2015 Initiative: design, implementation, and baseline findings. Popul Health Metr. 2015;13:3.

13. Athey. Identification and Inference in Nonlinear Difference-in-Differences Models. Econometrica [Internet]. Wiley Online Library; 2006 [cited 2018 Jan 18]. Available from: http://onlinelibrary.wiley.com/doi/10.1111/j.1468-0262.2006.00668.x/abstract

14. Dansereau E, McNellan CR, Gagnier MC, Desai SS, Haakenstad A, Johanns CK, et al. Coverage and timing of antenatal care among poor women in 6 Mesoamerican countries. BMC Pregnancy Childbirth. 2016;16:234.

15. Mokdad A. Supply-side interventions to improve health: findings from the Salud Mesoamérica Initiative. PLOS ONE. Forthcoming;

16. Till SR, Everetts D, Haas DM. Incentives for increasing prenatal care use by women in order to improve maternal and neonatal outcomes. Cochrane Database Syst Rev. 2015;CD009916.

17. Morris SS, Flores R, Olinto P, Medina JM. Monetary incentives in primary health care and effects on use and coverage of preventive health care interventions in rural Honduras: cluster randomised trial. The Lancet. 2004;364:2030–7.

18. Rwanda Ministry of Health, University of Rwanda, World Bank Group. Evaluation of the Rwanda Community Performance-Based Financing Program [Internet]. 2016 [cited 2018 Jul 17]. Available from: https://www.rbfhealth.org/event/evaluation-rwanda-community-performance-based-financing-program

19. Basinga P, Gertler PJ, Binagwaho A, Soucat ALB, Sturdy J, Vermeersch CMJ. Effect on maternal and child health services in Rwanda of payment to primary health-care providers for performance: an impact evaluation. Lancet Lond Engl. 2011;377:1421–8.

20. Perakis R, Savedoff W. Does Results-Based Aid Change Anything? Pecuniary Interests, Attention, Accountability and Discretion in Four Case Studies: Policy paper 052 [Internet]. Center for Global Development, Washington D.C.; 2015 Feb. Available from: https://www.cgdev.org/publication/does-results-based-aid-change-anything-pecuniary-interests-attention-accountability-and

21. Klingebiel S, Janus H. Results-Based Aid: Potential and Limits of an Innovative Modality in Development Cooperation. Int Dev Policy Rev Int Polit Dév [Internet]. 2014 [cited 2018 Jul 17];5. Available from: http://journals.openedition.org/poldev/1746

22. Paul E, Albert L, Bisala BN, Bodson O, Bonnet E, Bossyns P, et al. Performance-based financing in low-income and middle-income countries: isn't it time for a rethink? BMJ Glob Health. 2018;3:e000664.

23. Lim SS, Stein DB, Charrow A, Murray CJL. Tracking progress towards universal childhood immunisation and the impact of global initiatives: a systematic analysis of three-dose diphtheria, tetanus, and pertussis immunisation coverage. Lancet Lond Engl. 2008;372:2031–46.

24. Hartgerink C, Wicherts J, Assen M van. The value of statistical tools to detect data fabrication. Res Ideas Outcomes. 2016;2:e8860.

25. Phillips DL, Clancy KJ. Some Effects of "Social Desirability" in Survey Studies. Am J Sociol. 1972;77:921–40.

26. Hernandez B, Colombara DV, Gagnier MC, Desai SS, Haakenstad A, Johanns C, et al. Barriers and facilitators for institutional delivery among poor Mesoamerican women: a cross-sectional study. Health Policy Plan. 2017;32:769–80.

27. Rios-Zertuche D, Zúñiga-Brenes P, Palmisano E, Hernández B, Schaefer A, Johanns CK, et al. Methods to measure quality of care and quality indicators through health facility surveys in low- and middle-income countries. Int J Qual Health Care [Internet]. [cited 2018 Jul 25]; Available from: https://academic.oup.com/intqhc/advance-article/doi/10.1093/intqhc/mzy136/5039910

28. Heath I, Hippisley-Cox J, Smeeth L. Measuring performance and missing the point? BMJ. 2007;335:1075–6.

29. Bcheraoui CE, Palmisano EB, Dansereau E, Schaefer A, Woldeab A, Moradi-Lakeh M, et al. Healthy competition drives success in results-based aid: Lessons from the Salud Mesoamérica Initiative. PLOS ONE. 2017;12:e0187107.

Figure 1a: Month of first ANC visit at baseline and follow-up in SMI Intervention and control areas, as recorded in medical records from non-complicated pregnancies

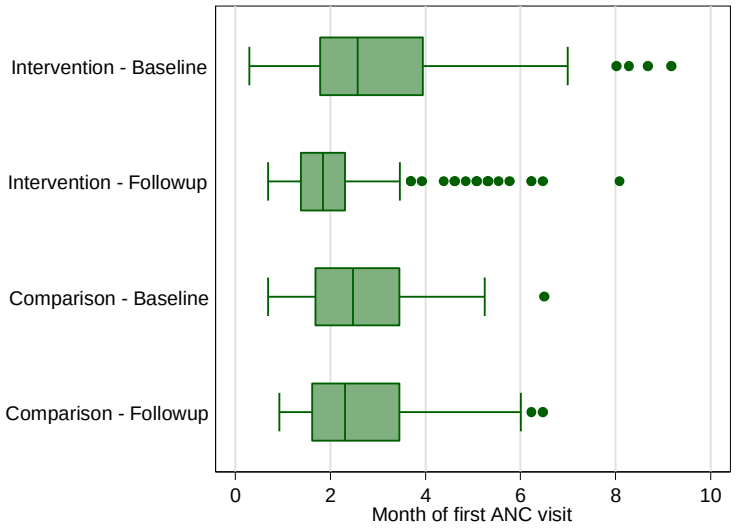


Figure 1b: Month of first skilled ANC visit during most recent pregnancy at baseline and follow-up in SMI Intervention and control areas, according to women's report in household surveys

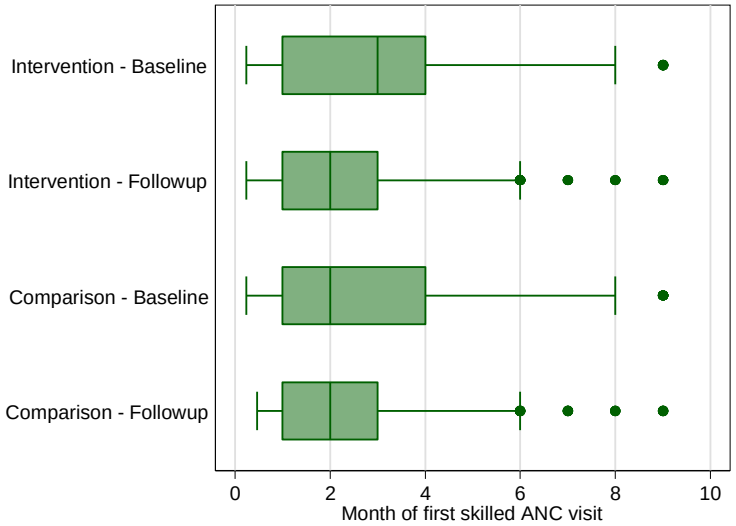


Figure 2a: Percent of women receiving standard diagnostic and preventive services at baseline and follow-up in SMI intervention and control areas, as recorded in medical records from non-complicated pregnancies

	Treatment			Comparison		
Medical record review	Baseline	Follow up	Change	Baseline	Follow up	Change
Blood glucose	30%	99%	69PP	48%	92%	44PP
HIV	94%	99%	5PP	88%	93%	5PP
Syphilis (VDRL or RPR*)	84%	99%	15PP	90%	91%	1PP
Anemia (Hb)	83%	99%	16PP	83%	68%	-15PP
Urinalysis	82%	99%	17PP	88%	91%	3PP
Blood group	86%	99%	13PP	84%	93%	9PP
Rh factor	86%	99%	13PP	84%	93%	9PP
At least 4 visits according to norm: weight, blood pressure, fetal movement, fundal height (if >=22 weeks), fetal heart rate (if >=20 weeks)	86%	95%	9PP	85%	82%	-2PP

*RPR only asked at follow-up

Figure 2b: Percent of women receiving standard diagnostic and preventive services at any point during most recent pregnancy at baseline and follow-up in SMI intervention and control areas, according to women's report in household surveys

	Treatment			Comparison		
Woman reported in household survey	Baseline	Follow up	Change	Baseline	Follow up	Change
Blood test	96%	99%	3PP	97%	96%	-1PP
Blood glucose	57%	78%	21PP	58%	78%	20PP
HIV	62%	79%	17PP	58%	82%	24PP
Syphilis	50%	66%	16PP	51%	61%	10PP
Anemia	82%	93%	11PP	82%	91%	9PP
Urine	75%	94%	19PP	84%	94%	10PP
Weight	98%	100%	2PP	99%	99%	0PP
Blood pressure	97%	99%	2PP	97%	99%	2PP
Fundal height	70%	96%	26PP	76%	96%	20PP
Fetal heartbeat	87%	97%	10PP	89%	98%	9PP

Table 1: Sample characteristics of ANC medical records and households surveyed in Honduras at baseline and followup in intervention and comparison areas

	SMI intervention areas		Comparison areas		
	Baseline	Follow-up	Baseline	Follow-up	Total
ANC medical records sample					
N	125	272	58	76	531
Age (median)	23	23	23	22.5	23
Married	91%	89%	87%	90%	89%
Literate	88%	97%	91%	97%	95%
Education					
None	4%	3%	7%	4%	4%
Primary	82%	87%	82%	84%	85%
Secondary	13%	10%	11%	12%	11%
University	1%	0%	0%	0%	0%
Household sample					
N	589	678	573	302	2142
Age (median)	24	26	25	26	25
Monthly expenditure (median per	121	115	161	197	143
Children (median)	2	2	2	2	2
Married	78%	85%	76%	75%	79%
Literate	64%	74%	67%	83%	72%
Education					
None	8%	11%	6%	5%	7%
Primary	73%	67%	69%	56%	66%
Secondary	9%	13%	10%	23%	14%
Baccalaureate/University	10%	10%	15%	17%	13%
Female head of household	18%	33%	21%	39%	28%
Mobile phone ownership	69%	69%	78%	79%	74%
Media exposure (past month)	81%	67%	79%	81%	77%
Prior obstetric complication	11%	12%	11%	9%	11%
CHW interaction	11%	15%	13%	6%	11%
Satisfied at last visit	97%	96%	93%	93%	95%
Bono 10,000 participation	21%	20%	23%	11%	19%

Table 2: Change in ANC indicators from baseline to followup in SMI intervention and comparison areas in Honduras, using ANC medical records and household data

	SMI Intervention Areas			Comparison Areas		
	Baseline	Followup	Change	Baseline	Followup	Change
ANC medical records						
Timeliness	59.2%	89.7%	30.5%	60.3%	69.7%	9.4%
Quality	23.2%	94.1%	70.9%	43.1%	60.5%	17.4%
Household data						
Timeliness	47.2%	63.0%	15.8%	55.7%	64.6%	8.9%
Quality	24.7%	48.1%	23.3%	22.0%	40.1%	18.0%
Utilization	87.8%	94.1%	6.3%	83.2%	80.1%	-3.1%

Table 3: Regression results from difference-in-differences analysis of SMI impact on ANC outcomes in Honduras

Medical record outcomes	**Main Model**	**Sensitivity Analyses**				
	Record-level adjusted	Record-level t-test	Facility-level t-test	Record-level clustered	Record-level among repeated facilities	Facility-level among repeated facilities
SMI effect sizes [a]						
Timeliness: ANC in 1st trimester	0.22*	0.21*	0.23	0.21*	0.02	-0.05
	[0.01, 0.42]	[0.05, 0.37]	[-0.00, 0.46]	[0.01, 0.41]	[-0.22, 0.26]	[-0.39, 0.28]
Quality: >=4 ANC, best practices	0.53***	0.53***	0.51***	0.53***	0.42	0.33
	[0.23, 0.84]	[0.39, 0.68]	[0.24, 0.77]	[0.23, 0.84]	[-0.03, 0.88]	[-0.07, 0.72]
Model specification						
N	531	531	120	531	208	54
Unit of analysis	Records	Records	Facilities	Records	Records	Facilities
Scope	All	All	All	All	Repeated facilities	Repeated facilities
SE clustered by facility	Yes	No	NA	Yes	Yes	NA
Adj for women's education	Yes	No	NA	No	Yes	NA

Household outcomes	**Main Model**	**Sensitivity Analyses**			
	Woman-level Adjusted [b]	Woma-level t-test	Municipality-level t-test	Woman-level Clustered	Matched women analysis
SMI effect sizes [a]					
Timeliness: ANC in 1st trimester	0.04	0.07	-0.03	0.03	0.01
	[-0.06, 0.14]	[-0.02,0.16]	[-0.20, 0.14]	[-0.07, 0.14]	[-0.15, 0.16]
Quality: >=4 ANC, best practices	0.04	0.05	-0.04	0.04	0.10
	[-0.11, 0.18]	[-0.03, 0.13]	[-0.26, 0.19]	[-0.11, 0.19]	[-0.08, 0.29]
Utilization: Any ANC at a public facility	0.05	0.09**	0.02	0.08	0.01
	[-0.03, 0.14]	[0.04, 0.15]	[-0.08, 0.12]	[-0.01, 0.17]	[-0.10, 0.13]
Model specification					
N	2142	2142	62	2142	1190
Unit of analysis	Women	Women	Municipalities	Women	Women
Scope	All	All	All	All	Matched women [c]
SE clustered by PSU	Yes	No	NA	Yes	Yes
Using sample weights	Yes	No	NA	Yes	Yes
Adjusting for covariates [b]	Yes	No	NA	No	Yes

Supplemental Box 1:

Sensitivity analysis:

ANC RECORDS DATA

- *Record-level t-test:* We used two-sided, two-sample t-tests of the null hypotheses that the change in average ANC timeliness and quality between baseline and follow-up was the same for intervention and comparison areas.
- *Facility-level t-test:* We calculated the average for each outcome by facility, and ran a two-sided, two-sample t-test at the facility level.
- *Record-level clustered analysis:* We accounted for the fact that women's records were clustered within health facilities.
- *Record-level adjusted for confounders* [MAIN ANALYSIS]: We built on the record-level clustered model by additionally controlling for potential confounders. Of the factors associated with ANC in a previously published analysis of SMI baseline data, the woman's education level was the only significant factor that was available in the medical record data.
- *Record-level adjusted among panel facilities:* We built on the record-level adjusted model by additionally restricting the sample of women's medical records to those coming from facilities we had sampled at both baseline and follow up (approximately half of facilities).
- *Facility-level panel analysis (medical records analysis):* Among facilities sampled at both baseline and follow-up, we calculated the average for each outcome by facility, and ran a two-sided paired t-test at the facility level.
- *Sub-group analysis:* For each of the above specifications, we also estimated the treatment effect within the following sub-groups: women with less education (primary or pre-primary); women with more education (secondary or beyond), to see if SMI had differential effects based on education status.

HOUSEHOLD SURVEY DATA

- *Woman-level t-test:* We used two-sided, two-sample t-tests of the null hypotheses that the change in ANC timeliness and quality between baseline and follow-up was the same for intervention and comparison areas.
- *Municipality-level t-test:* We calculated the average for each outcome by municipality (which was the unit of treatment assignment for SMI), and ran a two-sided, two-sample t-test at the municipality level.
- *Woman-level clustered analysis:* We accounted for the fact that women were clustered within survey segments. This specification also accounted for the sampling weights.
- *Woman-level adjusted for confounders* [MAIN ANALYSIS]: We built on the woman-level clustered model by additionally controlling for potential confounders.
- *Woman-level matched analysis:* The household sample was not designed to include repeat measurements of women, households, segments, or municipalities from baseline to follow-up. Therefore, we built on the woman-level adjusted model by using a novel matching method to create a 'quasi-panel' of women. We used Coarsened Exact Matching (CEM) to match women and achieve balance across the four samples (baseline intervention, baseline comparison, follow up intervention, and follow up comparison). We first matched baseline intervention

women with baseline comparison women based on the same characteristics we controlled for in the woman-level adjusted model. Women who did not achieve a match were trimmed from the dataset. Next, we matched the remaining baseline intervention women with follow-up intervention women, and the remaining baseline comparison women with follow-up comparison women. We trimmed the dataset to retain all matched women, as opposed to imposing 1:1 matches. We calculated each woman's inverse probability of inclusion following matching, and multiplied this by the survey sampling weights to create the final weights for analysis.

  - Following matching, the trimmed household dataset consisted of 1,190 women. As shown in Supplemental Table 1, matching achieved balance across the matching criteria, as well as most other characteristics, with the exception of literacy rates and community health worker interactions (which the baseline analysis did not find to be significantly associated with the outcomes). Compared to the un-trimmed sample, the matched group had higher literacy, greater education, more media exposure, and fewer prior obstetric complications.

- *Sub-group analyses:* For each of the above specifications, we also estimated the treatment effect within the following sub-groups: women in the poorest quintile (relative to other SMI households); women in the wealthiest quintile; women with less education (primary or pre-primary); women with more education (secondary or beyond), to see if SMI had differential effects based on wealth or education status.

Supplemental Table 1: Matched Samples

	SMI intervention areas		Comparison areas		
	Baseline	Follow-up	Baseline	Follow-up	Total
Medical records sample (After restricting to panel facilities)					
N	48	102	25	33	208
Age (median)	23.5	23.5	23	21	23
Married	92%	87%	83%	94%	89%
Literate	96%	98%	100%	100%	98%
Education*					
None	4%	2%	0%	0%	2%
Primary	85%	92%	92%	88%	90%
Secondary	10%	6%	8%	12%	8%
University	0%	0%	0%	0%	0%
Household sample (After matching)					
N	308	350	334	198	1190
Age (median)	24	24	25	26	25
Monthly expenditure (median per capita, USD)*	146	151	154	145	146

Children (median)*	2	2	2	2	2
Married	79%	84%	73%	76%	78%
Literate	73%	90%	73%	84%	80%
Education*					
None	1%	1%	1%	1%	1%
Primary	67%	67%	67%	67%	67%
Secondary	19%	19%	19%	19%	19%
Baccalaureate/University	13%	13%	13%	13%	13%
Female head of household	20%	32%	23%	35%	27%
Mobile phone ownership*	76%	76%	76%	76%	76%
Media exposure (past month)*	82%	82%	82%	82%	82%
Prior obstetric complication*	3%	3%	3%	3%	3%
CHW interaction	12%	15%	9%	8%	11%
Satisfied at last visit	97%	96%	95%	94%	96%
Bono 10,000 participation*	13%	19%	25%	9%	16%

Supplemental Table 2: Indicator values in SMI intervention and comparison areas at baseline and follow-up, in panel facilities and matched household sample

Medical records: From panel facilities

	Intervention			Comparison		
	Baseline	Followup	Change	Baseline	Followup	Change
Timeliness	66.7%	85.3%	18.6%	56.0%	72.7%	16.7%
Quality	10.4%	92.2%	81.7%	24.0%	63.6%	39.6%

Household data: Trimmed after matching

	Intervention			Comparison		
	Baseline	Followup	Change	Baseline	Followup	
Timeliness	42.3%	58.0%	15.7%	52.3%	66.7%	14.5%
Coverage	74.1%	86.3%	12.2%	76.2%	87.4%	11.2%
Quality	22.5%	39.3%	16.9%	26.1%	36.6%	10.5%

Supplemental Table 3: Sub-group and sensitivity analyses from DID analysis of the effect of SMI on ANC indicators in Honduras

Medical records: Women with primary or no education

	Record-level adjusted	Record-level t-test	Facility-level t-test	Record-level clustered	Record-level among panel facilities	Facility-level among panel facilities
SMI effect sizes [a]						
Timing: ANC in 1st trimester	0.20 [-0.03, 0.43]	0.20* [0.02, 0.37]	0.28* [0.20, 0.53]	0.20 [-0.03, 0.43]	0.03 [-0.22, 0.28]	0.02 [-0.35, 0.40]
Coverage: >=4 ANC	0.12 [-0.04, 0.27]	0.21 [-0.01, 0.24]	0.24* [0.06, 0.43]	0.12 [-0.04, 0.27]	-0.08 [-0.23, 0.06]	0.07 [-0.28, 0.42]
Quality: >=4 ANC, best practices	0.54** [0.21, 0.88]	0.54*** [0.39, 0.70]	0.59*** [0.31, 0.86]	0.54** [0.21, 0.88]	0.49 [-0.02, 0.99]	0.44 [-0.00, 0.88]
Model specification						
N	459	459	118	459	191	52
Unit of analysis	Records	Records	Facilities	Records	Records	Facilities
Scope	All	All	All	All	Panel facilities	Panel facilities
SE clustered by facility	Yes	No	NA	Yes	Yes	NA
Adjusting for women's education	Yes	No	NA	Yes	Yes	NA

Medical records: Women with secondary or university education

	Record-level adjusted	Record-level t-test	Facility-level t-test	Record-level clustered	Record-level among panel facilities	Facility-level among panel facilities
SMI effect sizes [a]						
Timing: ANC in 1st trimester	0.23 [-0.36, 0.83]	0.23 [-0.30, 0.77]	0.07 [-0.55, 0.70]	0.23 [-0.36, 0.83]	-0.27 [-1.09, 0.55]	-0.33 [-1.76, 1.09]
Coverage: >=4 ANC	0.08 [-0.36, 0.51]	0.08 [-0.24, 0.39]	-0.01 [-0.30, 0.29]	0.08 [-0.36, 0.51]	-0.30 [-1.62, 1.02]	-0.33 [-1.70, 1.04]
Quality: >=4 ANC, best practices	0.42 [-0.35, 1.18]	0.42 [-0.03, 0.87]	0.27 [-0.29, 0.84]	0.42 [-0.35, 1.18]	-0.17 [-0.76, 0.42]	-0.33 [-1.70, 1.04]
Model specification						
N	57	57	40	57	17	10
Unit of analysis	Records	Records	Facilities	Records	Records	Facilities
Scope	All	All	All	All	Panel facilities	Panel facilities
SE clustered by facility	Yes	No	NA	Yes	Yes	NA
Adjusting for women's education	Yes	No	NA	Yes	Yes	NA

Medical records: Impact of SMI on a continuous indicator of quality, measured as the percent of required tests/procedures performed during the pregnancy

	Record-level adjusted	Record-level t-test	Facility-level t-test	Record-level clustered	Record-level among panel facilities	Facility-level among panel facilities
SMI effect sizes [a]						
Quality (continuous variable): % of tests/procedures performed	0.13* [0.00, 0.26]	0.13** [0.05, 0.21]	0.16* [0.02, 0.30]	0.13* [0.00, 0.26]	0.08 [-0.11, 0.26]	0.12 [-0.07 0.31]
Model specification						
N	531	531	120	531	208	54
Unit of analysis	Records	Records	Facilities	Records	Records	Facilities
Scope	All	All	All	All	Panel facilities	Panel facilities
SE clustered by facility	Yes	No	NA	Yes	Yes	NA
Adjusting for women's education	Yes	No	NA	Yes	Yes	NA

[a] SMI impact is the coefficient for the interaction between the Follow-up period and Intervention facility dummies in regression models.

Household data: Women with primary or no education

	Woman-level Adjusted [b]	Woman-level t-test	Municipality-level t-test	Woman-level Clustered	Woman-level matched analysis
SMI effect sizes [a]					
Timing: ANC in 1st trimester	0.06 [-0.06, 0.18]	0.09 [-0.02, 0.19]	-0.05 [-0.27, 0.17]	0.06 [-0.06, 0.18]	0.06 [-0.11, 0.24]
Coverage: >=4 ANC	0.13* [0.02, 0.25]	0.11* [0.02, 0.20]	0.13 [-0.03, 0.28]	0.13* [0.02, 0.25]	0.05 [-0.11, 0.21]
Quality: >=4 ANC, best practices	0.10 [-0.05, 0.25]	0.09 [-0.01, 0.18]	-0.00 [-0.25, 0.25]	0.10 [-0.05, 0.25]	0.09 [-0.11, 0.28]
Model specification					
N	1576	1576	62	1576	867
Unit of analysis	Women	Women	Municipalities	Women	Women
Scope	All	All	All	All	Matched women [c]
SE clustered by PSU	Yes	No	NA	Yes	Yes
Using sample weights	Yes	No	NA	Yes	Yes
Adjusting for covariates [b]	Yes	No	NA	Yes	Yes

Household data: Women with secondary, baccalaureate or university education

	Woman-level Adjusted [b]	Woman-level t-test	Municipality-level t-test	Woman-level Clustered	Woman-level matched analysis
SMI effect sizes [a]					

Timing: ANC in 1st trimester	0.08 [-0.12, 0.27]	0.16 [-0.00, 0.32]	0.06 [-0.18, 0.30]	0.08 [-0.12, 0.27]	-0.04 [-0.33, 0.26]
Coverage: >=4 ANC	-0.05 [-0.20, 0.09]	-0.03 [-0.15, 0.08]	-0.07 [-0.22, 0.08]	-0.05 [-0.20, 0.09]	-0.13 [-0.33, 0.08]
Quality: >=4 ANC, best practices	-0.06 [-0.29, 0.17]	-0.01 [-0.17, 0.15]	0.00 [-0.33, 0.33]	-0.06 [-0.29, 0.17]	0.05 [-0.25, 0.34]
Model specification					
N	566	566	58	566	323
Unit of analysis	Women	Women	Municipalities	Women	Women
Scope	All	All	All	All	Matched women [c]
SE clustered by PSU	Yes	No	NA	Yes	Yes
Using sample weights	Yes	No	NA	Yes	Yes
Adjusting for covariates [b]	Yes	No	NA	Yes	Yes

Household data: Women in the lowest household expenditure quintile (relative to the SMI sample)

	Woman-level Adjusted [b]	Woman-level t-test	Municipality-level t-test	Woman-level Clustered	Woman-level matched analysis
SMI effect sizes [a]					
Timing: ANC in 1st trimester	-0.10 [-0.32, 0.12]	-0.02 [-0.26, 0.22]	0.06 [-0.26, 0.37]	-0.05 [-0.29, 0.18]	-0.21 [-0.56, 0.15]
Coverage: >=4 ANC	0.26** [0.10, 0.43]	0.25*** [0.14, 0.37]	0.09 [-0.03, 0.21]	0.25** [0.08, 0.43]	0.11* [0.01, 0.21]
Quality: >=4 ANC, best practices	-0.10 [-0.42, 0.22]	-0.11 [-0.33, 0.11]	-0.17 [-0.53, 0.19]	-0.07 [-0.39, 0.24]	-0.05 [-0.51, 0.40]
Model specification					
N	421	421	56	421	164
Unit of analysis	Women	Women	Municipalities	Women	Women
Scope	All	All	All	All	Matched women [c]
SE clustered by PSU	Yes	No	NA	Yes	Yes
Using sample weights	Yes	No	NA	Yes	Yes
Adjusting for covariates [b]	Yes	No	NA	Yes	Yes

Household data: Impact of SMI on a continuous indicator of quality, measured as the percent of required tests/procedures performed during the pregnancy

	Woman-level Adjusted [b]	Woman-level t-test	Municipality-level t-test	Woman-level Clustered	Woman-level matched analysis
SMI effect sizes [a]					
Quality (continuous variable): % of tests/procedures performed	0.03 [-0.03, 0.09]	0.02 [-0.01, 0.06]	-0.00 [-0.08, 0.08]	0.02 [-0.04, 0.08]	0.02 [-0.06, 0.09]
Model specification					
N	2142	2142	62	2142	1190
Unit of analysis	Women	Municipalities	Women	Women	Women
Scope	All	All	All	All	Matched women [c]
SE clustered by PSU	No	NA	Yes	Yes	Yes
Using sample weights	No	NA	Yes	Yes	Yes
Adjusting for covariates [b]	No	NA	Yes	Yes	Yes

[a] SMI impact is the coefficient for the interaction between the Follow-up period and Intervention area dummies in regression models.
[b] Covariates include: household wealth quintile as compared to other SMI households, number of children, education level, mobile phone ownership, media exposure in the past month, and previous obstetric complications.
[c] Women were matched based on the same list of covariate variables

Supplemental Figure 1: Week of first ANC visit recorded in medical records and self-reported by women