Health Insurance and Child Health: Evidence from Seguro Popular

Gina Turrini

Duke University

Gabriela Farfán

World Bank

María Genoni

World Bank

Duncan Thomas

Duke University

Andrea Velasquez

University of Colorado Denver

Preliminary version - September 2015

Public health insurance programs are being expanded across the globe despite limited evidence regarding their impacts on population health. Exploiting the roll-out of Seguro Popular, a large-scale program that provides public health insurance to about half Mexico's population, this research isolates the causal impact of the program on child health and nutrition measured by height-for-age. Drawing on insights from the biology of human linear growth during the first few years of life, we use rich longitudinal population-representative data, the Mexican Family Life Survey, in combination with administrative program data and establish that Seguro Popular has had, at best, a modest impact on child nutritional status. Program effects in a community are larger after the program has been established for several years, suggesting supply-side factors may be critical impediments to fulfilling program goals. The results have important implications for the design, roll-out, and evaluation of public health programs.

INTRODUCTION

In 2002, the Mexican government began to roll out a large-scale public health insurance program, Seguro Popular, in an attempt to address inequalities in access to health care across the country. Although many countries in Latin American and elsewhere have similar programs, Seguro Popular is unusual in terms of its size and scope (Atun et al. 2015). It covers a very large number of interventions, medications, and treatments, and nearly half the population is eligible. The program is so large and required so much money, time, and investment that it took nearly a decade to implement.

This paper provides new evidence on the impact of the program on an important indicator of population health: child height given age and gender. The lessons from Mexico's Seguro Popular might well be informative for the other middle-income countries that implement similar programs. Despite this, there has been very little research on the impact of the program on health, and particularly on child health. Most of the previous research on Seguro Popular has been on the impact of the program on health insurance coverage, use of health care services, and household spending on health care. While these outcomes are significant, it is also important to measure the impact of this large public insurance program on health, particularly in a population that still faces a large burden of disease.

Child height is a useful and important measure of child health for a number of reasons. We know, for instance, that malnutrition in childhood is associated with a number of worse later life outcomes, including lower adult earnings, worse cognitive functioning, and worse adult health status (Strauss and Thomas 1995). In addition, child height has a unique feature that makes it useful for empirical estimation. The nutrition and human biology literatures have demonstrated that height is sensitive to nutritional inputs and disease insults during the first few years of life – often referred to as the critical first 1,000 days. After that point, absent a large shock to nutrition or the health environment, the child's linear growth trajectory is thought to be determined, meaning that attained stature as an adult is also largely determined. During the critical early period, child height can be

2

affected through the quantity and quality of nutrition and through the impacts of some childhood diseases like diarrhea (for examples of other papers making use of the biology of child height, see Frankenberg, Suriastini and Thomas 2005; Duflo 2000; Duflo 1999; and Farfan et al. 2013). Seguro Popular is designed to assure access to health-related services that potentially affect both of these channels.

This analysis uses the insights from the biology of human growth, together with the gradual roll-out of the program, to identify the effect of Seguro Popular on child height and, thereby, on child health. This project uses a panel data set (the Mexican Family Life Survey, described in more detail below) which spans most of the roll-out of the program and also includes extensive information on health status (including anthropometric measures such as child height) and utilization, insurance coverage, and socioeconomic status. More complete descriptions of the important features of the program and of the data used for this work are in the next two sections. This is followed by a discussion of the identification challenges faced in estimating the program's impact along with an explanation of the estimation strategy. Next is a presentation of results, followed by discussion and conclusion. Overall, we find some limited evidence of an effect of the program on child height, but only when we compare children who lived in municipalities that got the program in 2002 with children who lived in municipalities that got the program after 2008. The results also suggest that any effect may have taken quite a while to emerge.

SEGURO POPULAR

Prior to the introduction of Seguro Popular in 2002, the vast majority of Mexican families who were covered by health insurance were employed either in the formal sector by private firms or in the public sector. Workers in the formal sector are covered by benefits provided by the Mexican Institute for Social Security (IMSS) while public sector workers are covered by a similar program run by the Institute of Social Security and Services for Civil Servants (ISSSTE). By 2000, approximately 40% of the

3

Mexican population was covered by IMSS, 7% was covered by IMSSTE, and 3% was covered by private insurance (Knox 2008). As a result, half of the Mexican population (more than 50 million people) remained without any form of health insurance (Frenk 2006). This prompted the creation of Seguro Popular, a government-implemented insurance program available for all Mexican households who do not otherwise have health insurance. The program has several main goals, including protecting individuals and families from catastrophic healthcare spending, lowering out-of-pocket healthcare expenditures, and ensuring access to accredited healthcare providers and facilities (King et al. 2007)

The costs of the program are primarily shared between the state and federal governments, although a small amount is contributed by annual fees from enrollees. The state governments administer the program and decide how to allocate funds. However, they are required to ensure that there are enough accredited service providers and infrastructure so that people who affiliate actually have access to the services guaranteed. As a result, a significant amount of initial program funds were allocated to investment in health infrastructure, although some commentators suggest that there were areas, particularly in rural municipalities, where the minimum requirements were not always initially met (Pfutze 2014; Barros 2008).

Enrollees in Seguro Popular have access to a large number of treatments, medical services, and medicines without co-pay. Upon enrollment, households or individuals are assigned to a primary-care physician and public clinic (Conti and Ginja 2014). The number of interventions and medicines covered has grown as the program has rolled out. By 2006, Seguro Popular covered the treatments for diseases responsible for 95% of the burden of disease in the country (King et al. 2007). The program encompasses four main components. The first and largest is Catálago Universal de Servicios de Salud (CAUSES) which has been in place since the start of Seguro Popular and is operated at the state level. In 2004, CAUSES covered 91 interventions and 142 drugs; by 2012 it had expanded to 284 interventions and approximately 400 drugs. The second component, Fondo de Protección contra Gastos Catastróficos

(FPGC), has also been in place since the program began but is managed at the federal level. FPGC is primarily a trust fund to cover catastrophic events (such as cancer treatment or surgery), and by 2012 included 58 interventions (Pfutze 2014).

There were two modules added to Seguro Popular after the program began. Seguro Médico para una Nueva Generación (SMNG) began December 1, 2006 and focuses on infant and child health. Every child born after SMNG began and their immediate families (if they are not affiliated with any social security insurance through the government) are now automatically covered by Seguro Popular. SMNG covers children and their families until the children are five years old (National Commission of Social Health Protection). This module covers 131 interventions that focus on early childhood health. The primary goals of SMNG were to lower infant and child disease and mortality while preventing catastrophic expenditures for families of young children. Seguro Popular and SMNG both include interventions and treatments for young children, while Seguro Popular also includes some interventions for pregnant women. SMNG includes treatment for a very wide range of conditions, including genetic issues, complications from birth, nutrition and metabolic disorders, hearing and vision problems, and other potentially high-cost health issues. Children covered by SMNG are also covered by Seguro Popular, which includes the treatments for the most common childhood diseases, a complete physical after birth, infant and child vaccinations, neonatal screenings, and training for mothers on breastfeeding and nutrition (CAUSES 2012). They are also covered for 44 interventions (for high-cost, low-likelihood illnesses) focusing specifically on young children through FP (Rodríguez-Ortega 2012; Muñoz-Hernández 2012). Embarazo Saludable (ES), added in May 2008, does not cover new interventions but simply extends automatic coverage to pregnant women and their families if they do not have insurance coverage (Pfutze 2014).

There are a number of services covered under Seguro Popular that could be possible pathways for the program to affect child nutrition. For instance, the program covers a range of interventions

relating to childbirth, neonatal care, and infant preventive health care like vaccinations and screenings. The program also covers education on breastfeeding and infant nutrition, which could very directly affect child nutrition. Similarly, for children who are older than infants but younger than five years of age the program covers complete health check-ups, which include measurement of height and weight and training of parents on nutrition and feeding. Childhood diarrhea is another possible mechanism that could affect nutrition and height. Seguro Popular includes treatment for diarrhea, including oral rehydration and training for parents on how to prevent and recognize diarrhea (CAUSES 2014). More indirectly, if SP lowers out-of-pocket expenditure for health care more generally and reduces the chances of catastrophic healthcare spending, this could also potentially impact child nutrition simply by freeing up resources within a household.

To be covered by Seguro Popular, individuals or households must actively affiliate with the program, with a few exceptions. In addition to the automatic coverage of pregnant women, infants, and their families added to the program later and discussed above, individuals already enrolled in Oportunidades are affiliated with Seguro Popular automatically. In general, however, people must choose to enroll and, as part of that process, are required to pay an annual fee based on self-reported income. The lowest two income quintiles do not have to pay anything (Barros 2008). As a result of income being self-reported, although there is technically a fee for households with higher incomes, the Department of Health estimated that the vast majority of households (more than 99%) do not pay for coverage (Pfutze 2014).

The program began with a pilot in five states in 2002 and continued with full roll-out beginning in 2004. There were, however, 15 states which began roll-out before there was an official agreement with the federal government in 2002 and 2003 (Bosch et al. 2012; Conti and Ginja 2014). Due to financial and logistical constraints, the program was to be implemented gradually over seven years, with an annual cap on new enrollees. Each year, the state and federal governments negotiated the target

6

number of households the state would enroll in each quarter of the following year until states reached

their official final target number (theoretically equal to the number of uninsured households in their

state). Officially, states with low insurance coverage, high numbers of uninsured individuals in the first

six income deciles, the ability to provide services, high demand for enrollment, and a sufficient budget

were to be targeted first (Conti and Ginja 2014). Within states, Seguro Popular was rolled out at the

municipality level, meaning that a municipality could begin affiliating enrollees once their health care

facilities met minimum infrastructure and human resource requirements (Knox 2008).

While officially Seguro Popular was intended to be rolled out first to poorer states and states

with high numbers of uninsured households, there is disagreement about whether this happened in

practice. Some observers have suggested that politics played a role in determining when and how

quickly the program rolled out to states and municipalities. For instance, Barros (2008) argues that

smaller states and those affiliated with the governing party tended to have the program implemented

sooner. His explanation of this phenomenon is that for smaller states, it was easier to achieve full

coverage, and these states could be used by the federal government as examples of the success of the

program. Knox (2008) notes that the five states included in the pilot program had health ministers who

were affiliated with the Minister of Health. Officially, these five pilot states were chosen because they

had the infrastructure to provide the guaranteed services, they had large semi-urban populations, and

because they had a large number of households who participated in social programs such as

Oportunidades (Conti and Ginja 2014). Although the Inter-American Development Bank (2012)

concluded that program roll-out was essentially random, it is not clear that this is in fact the case.

To get at this issue, we can look at whether date of roll-out seems to be related to municipality

resources for the municipalities in our sample. Table 1 (which uses data described in more detail below)

shows municipality averages for a select number of relevant characteristics at three dates, stratified by

when Seguro Popular was rolled out. For instance, the first two panels show measures of household

resources, including average education of the household head and average per capita expenditures. Generally, there is little evidence that Seguro Popular was rolled out first to municipalities that were poorer. If anything, in our data the opposite seems generally true. For instance, the average education (in years) of the household head was 6.3 for communities that got Seguro Popular in 2002 versus 4.9 for municipalities that got the program between 2006 and 2008 (and this difference is statistically significant). Interestingly, household heads in municipalities that got Seguro Popular after 2008 have higher average education, at 8.1 years, which is significantly higher than the average for any of the other three groups. There is no clear pattern in terms of the log of per capita expenditure, contrary to the government's stated policy of introducing Seguro Popular into poorer areas first. This may reflect the fact that roll-out was conditional on municipalities having the infrastructure to provide the required services.

Panels C and D show some of the municipality characteristics that were officially supposed to influence roll-out, including the fraction of households informally employed and the fraction of households without medical insurance. Similar to what was found for household characteristics it seems that if anything municipalities that had higher rates of formal employment and insurance coverage were targeted first. This does not align with the official stated policy of the government but is generally consistent with previous empirical evidence on the roll-out of the program.

DATA

The primary data source for this project will be the Mexican Family Life Survey (MxFLS), a panel survey launched in 2002. The first wave (MxFLS1) was conducted in 2002, the second (MxFLS2) in 2005-2006, and the third (MxFLS3) in 2009-2012. The baseline sample was representative at the regional, rural, urban, and national levels. It contained a sample of 35,677 individuals in 8,440 households, located in 150 communities spread over 16 states. During the second and third rounds, attempts were made to relocate all individuals from the first round even if they had migrated within Mexico or to the

United States. The attempts to follow and find households who had moved were quite successful; MxFLS2 and MxFLS3 found and re-interviewed nearly 90% of the original sample (for more information on MxFLS, see Rubalcava and Teruel 2006, Rubalcava and Teruel 2008, and Rubalcava and Teruel 2013).

MxFLS contains a rich variety of demographic and economic measures. Critical for this project is that an array of health and anthropometric information was collected from all household members. This includes information on height (or length for young infants) and weight, which were used to construct height-for-age z-scores. The survey also included extensive information on other health-related outcomes, including self-reported health, illness, use of health care (such as prenatal care), healthcare expenditures, and vaccinations. There is also information on household members, including education, income, and employment. For each household member at least 14 years old, the survey collects information on whether household members received benefits from social programs (such as Oportunidades) and if they are covered by health insurance (and if so, from what source). There were also surveys of municipality facilities, including health care facilities, which can provide some information on the availability of health care services within a municipality.

The data from MxFLS are supplemented with administrative data on the date that Seguro Popular was initially rolled-out within each municipality. This allows the creation of a measure, for each child, of how much exposure they had to the program during their early childhood based on when Seguro Popular was introduced into their community.

MEASURING SEGURO POPULAR'S EFFECT ON CHILD HEIGHT

Isolating the causal effect of Seguro Popular is complicated by a number of identification challenges, only some of which the previous literature has been able to address. While a few studies find positive health effects of the program (Sosa-Rubí et al. 2009; Aguilera and Marrufo 2006; Pfutze 2014; Bleich et al. 2007), others find no effect at all (Barros 2008; King et al. 2009). It is difficult to compare these results since they focus on different outcomes and use a range of methodologies.

Unfortunately, for the health outcomes these previous papers focused on, the effect of the program cannot be identified using observational data without making additional assumptions. For instance, several papers use the date of roll-out as an instrument for enrollment and then look at the relationship between enrollment and health outcomes. However, roll-out also coincided with an increase in spending on health infrastructure that may have affected families who were not enrolled as well. Therefore, the exclusion restriction is almost certainly violated. Similarly, a number of the papers assume that the date of roll-out was essentially random. If roll-out is related to the health environment, however, this will cause problems for their estimation. While one paper does use experimental data (King et al. 2009), it suffers from external validity concerns (the experiment was run in a small number of carefully selected municipalities which are very unlikely to be representative of the entire country) and only covers ten months (arguably not enough time for many health outcomes to respond to roll-out of health insurance). Reliance on additional and likely unrealistic assumptions is a shortcoming of the previous literature on Seguro Popular.

Looking at how the prevalence of various health outcomes changed over time can help to illustrate how analysis depending on random roll-out or other additional assumptions might lead to incorrect conclusions. Figure 1 shows, as an example, the percent of children aged 0 to 2 who had a cough in the previous four weeks in 2002, 2005, and 2009. The children are stratified based on when their municipality got Seguro Popular. Looking at children in municipalities with the program in 2002, there is a dramatic decline in reported cases of cough between 2002 and 2005. It might be tempting to conclude that this is the impact of the program. However, there is a very similar improvement between 2002 and 2005 for children who got the program between 2006 and 2008, although it could not have been caused by the program. Something else seems to have been going on as well. It is possible that the difference in improvement (a difference-in-difference) is significant if the areas that got Seguro Popular improved more. However, attributing that additional improvement to Seguro Popular requires

10

a parallel trends assumption that may not be reasonable if areas that got Seguro Popular earlier were on a different trajectory even before receiving the program.

This analysis improves on previous work first by using an identification strategy focusing on child height that has been successfully implemented in several similar projects (for instance, see Duflo 2001; Frankenberg, Suriastini and Thomas 2005; and Farfan et al. 2013). Child height is particularly well-suited for this type of analysis because exposure to Seguro Popular should only affect a child's height-for-age z-score if they are exposed between birth and approximately age four. Our definition of treatment, described in more detail below, will rely on this fact about the biology of child height.

When measuring the effect of the program, one cannot simply compare the height of children in households enrolled in Seguro Popular with children in households that have not enrolled. This comparison would be difficult to interpret because it does not take into account differences between households who do and do not choose to enroll in the program. The households who opt in are likely to be different than households who do not in ways that are not observed (for instance, they may have greater need for health insurance because of pre-existing health conditions). The primary analysis in this project will therefore be intent-to-treat, where treatment status will be a function of age of the child and of when Seguro Popular was introduced to their community, regardless of whether their family reports actually enrolling in the program. The intent-to-treat analysis can be supplemented by running similar regressions with households who, at baseline, were likely to be eligible because they worked in the informal sector or reported not having health insurance.

The comparison of children who were likely to benefit from Seguro Popular with those unlikely to benefit are still complicated by the non-random roll-out of the program. For instance, the official policy was that Seguro Popular be introduced in areas with higher levels of poverty first. If that policy was actually implemented, children in areas that got Seguro Popular right away might initially be at a disadvantage compared to children in areas that got Seguro Popular later, making it appear that the

program is having little to no effect when that is not in fact the case. If the observable characteristics

that were used to plan roll-out were known, it might be possible to control for those characteristics in

an analysis. However, the existing literature suggests that the roll-out process was far from transparent,

so even if the model includes a wide range of municipality-specific characteristics, there might still be

unobservable differences that cannot be accounted for.

Unfortunately, much of the previous literature relies on the roll-out being random for

identification. As previously discussed, it is not clear what municipality characteristics determined roll-

out, in addition to the ones that were officially stated. Figure 2 suggests that municipalities that got the

program at different points in time may have had very different health environments in 2002 when the

program began. Comparing children aged 0 to 2 in 2002 who got the program in their municipality

earliest with those that got it in their municipality latest we can see there is a large difference in the

reported rates of diarrhea for children before the program was implemented. This suggests that the

date of roll-out was not random and that there is reason to be concerned that date of roll-out is

correlated with the health status of children.

To address this concern, we will construct groups of municipalities that are likely to be fairly

similar (based on their date of roll-out) and look at the impact of the program within these municipality-

groups. The first group will consist of communities in which Seguro Popular was rolled out at some

point in 2002. The municipalities in the second group had Seguro Popular introduced between 2003 and

2005 (between the first and second waves of MxFLS). The third group consists of municipalities which

got Seguro Popular between 2006 and 2009 (between the second and third waves of MxFLS).

Municipalities where the program had not been rolled out by MxFLS3 in 2009 can be used as a reference

group.

Finally, the use of panel data helps in a few ways. First, panel data allows us to control for

household, municipality, and state characteristics. In addition, panel data helps in the interpretation of

height-for-age z-score. It has been found that in low and middle-income countries, height-for-age tends

to decline until weaning before recovering slightly and then remaining fairly constant (Martorell and

Habicht 1986). This makes it more challenging to interpret any comparisons of younger and older

children. Controlling for exact age in the analysis can partially account for these dynamics of height-for-

age over childhood. Using the three waves of MxFLS can strengthen the interpretation of the estimates

by comparing children in different communities measured at approximately the same age at different

points in time.

To this end, children in our sample will be separated into four bands based on their ages during

the different rounds of MxFLS. This will allow the separation of children based on whether they are at

an age where nutritional interventions should affect their height. The oldest children in the analysis are

nine years old. The upper age limit is imposed so that none of the children in the analysis should have

hit puberty, which causes additional complications in terms of the dynamics of growth. At each round of

MxFLS, children are divided into those who are *very young* (between 0 and 2), *young* (between 3 and 4),

old (between 5 and 6) and those who are *very old* (between 7 and 9).

The relationship between exposure and age is key for the analysis. The expected exposure for

the various age groups, which depends on whether their community got Seguro Popular while they were

in the critical window (0 to 4 years old), is outlined in Figure 3. For instance, there should be essentially

no effect of the program on the height of children when they are observed in 2002 (since they will have

experienced at most only a very short period of exposure to the program) and they are therefore

labeled in the figure as having experienced no impact of the program. However, the story is different

for children who were less than three years old in 2005 who lived in a municipality that got the program

in 2002 (the cell in the row labeled "2005: 0 to 2 years old" and the first column). They would have lived

their whole lives with the program in place and there should be a full effect on them. Slightly older

children in those same locations, such as those who are between 3 and 4 years old in 2005 (the cell in

the row labeled "2005: 3 to 4 years old" and first column), would have experienced only a partial effect of the program since it would have arrived to their municipality after they were born but still during their critical early years. They are labeled as experiencing some effect of the program. In the fourth column are children who live in municipalities that got Seguro Popular after 2009. For those children there should be no effect of the program at any of the three waves, and they will serve as a useful control group.

SAMPLE AND METHODS

Description of the sample population

A selection of descriptive household and child characteristics is presented in Table 2. In the first panel, height-for-age z-scores (calculated using World Health Organization references) are shown for each wave of the survey, split into two age groups and into four municipality-groups based on when their municipality got Seguro Popular (these divisions will be described in more detail with the empirical specification). In almost every column, there is a decline in height-for-age moving from the children who got the program in 2002 to those who got it between 2005 and 2008, but then height-for-age goes up for children who got the program after 2008. This is suggestive of worse off communities tending to receive the program later as we move from 2002 to 2008, contrary to the official policy but consistent with the literature on the roll-out. There is also a general increase over time in height-for-age in most of the subgroups.   The other household and child characteristics (in the second panel) are relatively stable over time, with the average child in the sample being 4.5 years old in a household with roughly 6 members. The average education level of the household head is about 6.5 years and average monthly log per capita expenditure is about 6.7 dollars.

Estimation

Given the definitions of age bands and municipality groups outlined above, we turn to the regression equation:

$$\theta_{icvt} = \alpha_{ct}^4 + \alpha_{ct}^1 I_1 + \alpha_{ct}^2 I_2 + \alpha_{ct}^3 I_3 + \gamma_c' X_{ivt} + \gamma_v + \varepsilon_{ivt} \qquad (1)$$

where θ_{ictv} is the height-for-age z-score of child *i*, in age band *c*, in municipality-group *v*, at time *t*. The superscripts on the α terms represent the municipality-groups (group 1 through group 4). The α_{ct}^4 terms are twelve different age group-time intercepts. There are intercepts for very young, young, old, and very old at time 0 (2002), intercepts for very young, young, old, and very old at time 1 (2005), and intercepts for very young, young, old, and very old at time 2 (2009). These coefficients estimate the height-for-age z-scores of children in group 4, who form the control (because they did not have Seguro Popular at all before 2009). The next three terms are interactions between age group-time intercepts and indicator variables for belonging to one of the three groups that did get Seguro Popular before MxFLS3. For instance, the α_{ct}^1 terms represents the interaction of each of the age group-time intercepts with an indicator for belonging to group 1. The subscripts for the age groups are A for very old, B for old, C for young, and D for very young. As a specific example, α_{C0}^2 represents the difference in average height-for-age at time 0 (MxFLS1) between young children who lived in community-group 2 (who got Seguro Popular between 2003 and 2005) and young children who lived in control communities (group 4 communities that did not get Seguro Popular until after MxFLS3). Each of the interaction terms can be interpreted in this fashion. The X covariates include individual gender and age in months. Finally, the model includes municipality fixed effects to help deal with municipality-level characteristics that might have affected roll-out. It is important to note these are municipality fixed effects, not municipality-group fixed effects (which would introduce issues of multicollinearity). The same municipality of residence from the first round is maintained throughout in case there is any selective migration to areas that got Seguro Popular earlier. Standard errors are clustered at the municipality-time-age group level

(clustering at the municipality, municipality-time, or municipality-age group level produces very similar results).

The first version of equation (1) includes all children who are nine or younger during the three waves. This is the truly intent-to-treat version of the model. These estimates can be compared with coefficients from the same regression limited to the sample most likely to benefit from the program. In one version, this will be children who live in households where no one works in the formal sector at baseline (since formal employees have health insurance by law). The second version will limit the sample to children who live in households where none of the adults report having health insurance at baseline. Comparing these estimates can give some suggestive evidence about whether the estimates from the intent-to-treat are being driven downward by including children whose families cannot benefit from the program.

Difference-in-difference estimates

The coefficients in equation (1) cannot simply be read as the effect of the program on child height because roll-out was not random. A modified version of a difference-in-difference will be used instead to get estimates of the causal impact of the program. There are a number of variations of the difference-in-difference that could be estimated; a few possible examples are outlined below.

The first comparison is between children of different ages in the same community group at the same time, which uses the variation in exposure which is the result of the critical window for nutrition and height. The gradual roll-out of the program allows the comparison of younger children who were likely to be fully or partially exposed with older children who were unexposed. As a specific example, $\alpha_{D1}^1 - \alpha_{A1}^1$ provides an estimate of the effect of full exposure by comparing (in 2005) very young children with very old children within communities that got the program in 2002 (group 1). This approach has the advantage that it sweeps out any characteristics common to municipalities within the same community group (which may have influenced when the program was rolled out). However, one

16

possible concern is with life-cycle dynamics that can cause younger and older children to have different height-for-age z-scores even in the absence of any program effect. Previous literature suggests that, particularly in lower and middle-income contexts, height-for-age z-scores tend to decline until weaning, increase slightly, and then stabilize. This can affect the comparison of younger and older children regardless of program impact. In addition, any time-varying community characteristics or time-trends that affected the height of older children differently than the height of younger children may also influence this estimate.

This suggests a second comparison that could provide an alternative estimate of the program effect. Instead of comparing children of different ages within a community group at the same point in time, children of the same age can be compared at different points in time. For instance, $\alpha_{D1}^{1} - \alpha_{D0}^{1}$ gives an estimate of the effect of full exposure by comparing children who were very young in 2005 with children who were very young in 2002 (all within communities that got the program in 2002). This addresses the issues of life-cycle dynamics but this comparison raises the concern of confounding time effects. If any changes occurred within these communities between 2002 and 2005 that affected child height besides the introduction of Seguro Popular, that could bias this estimate.

One way to look at whether time effects are influencing the estimate is to do another, similar comparison. Instead of comparing children who were likely exposed to those unlikely to be exposed, however, this comparison will look at older children who were all unlikely to have experienced an effect of the program. For instance, $\alpha_{A1}^{1} - \alpha_{A0}^{1}$ is the difference in height-for-age comparing very old children in 2005 with very old children in 2002, all within communities that got the program in 2002. None of these children should have been exposed to the program, so this difference can give an idea of time effects that were operating in the background that might be influencing the estimate of the program effect.

By combining these comparisons into various difference-in-difference estimates, a number of estimates of the program impact can be constructed. Depending on which combination of differences is used, these estimates will be unbiased under different assumptions and robust to different potential confounding factors. To get an estimate of the impact of the program on children who were fully exposed, for instance, we can calculate $(\alpha_{D1}^1 - \alpha_{A1}^1) - (\alpha_{D0}^1 - \alpha_{A0}^1)$. The first term gives an estimate of the program impact if there is no age effect. Subtraction of the second term gets rid of any age effect. However, for this to be capturing only the effect of the program, it needs to be true that there were no time effects that influenced the difference between very young and very old children in a different way in 2002 versus 2005. Looking at the same comparison for children who lived in communities that got the program after 2005, who should not have felt any impact of the program by MxFLS2, can give us an idea if this is an issue.

Another version of the difference-in-difference is $(\alpha_{D1}^1 - \alpha_{A1}^1) - (\alpha_{D1}^4 - \alpha_{A1}^4)$. The first term is the same as before, but now the second term looks at children in 2005 living in communities that got Seguro Popular after 2009 (a similar comparison is possible with children whose communities got the program between 2005 and 2008). This has the benefit that if time trends were the same in the two types of communities, this difference eliminates any age effects and time trends that affected older versus younger children differently. However, for this estimate to be unbiased it must be true that in the absence of the program the difference between very young and very old children would have been the same in 2005 in these two types of communities. This might not be true if there are different trends over time in these two locations.

A final type of difference-in-difference constructs estimates of full versus partial exposure. In particular, we can compare whether this estimate changes depending on when the program is rolled out. This may provide a sense of whether the impact of the program was changing over time (if, say, the program became more effective as it became more established or well-known). The effect of full versus

partial exposure can be estimated by calculating $(\alpha^1_{D2} - \alpha^1_{A2}) - (\alpha^4_{D2} - \alpha^4_{A2})$. This estimate can then be

compared with $(\alpha^2_{D2} - \alpha^2_{A2}) - (\alpha^4_{D2} - \alpha^4_{A2})$, which should also provide an estimate of the effect of full

versus partial exposure. If these are significantly different, it might suggest that the effect of the

program changed depending on when it was rolled out.

RESULTS

The results of estimating model (1) for the full sample are presented below in Table 3. The first

column gives the differences in height-for-age z-scores for children in communities that got Seguro

Popular in 2002 compared to children whose communities got the program after 2009 (the reference

group in the fourth column). The second and third columns compare the height-for-age of children in

the reference group to children in communities who got Seguro Popular in 2003-2005 and 2006-2008,

respectively. The model was estimated using ordinary least squares and standard errors (robust and

clustered at the municipality-year-cohort level) are reported below. Additional covariates that are not

shown include gender and age in months. In addition, the model includes a community fixed effect.

The coefficients in bolded blue are those that should reflect full exposure to the program. The

coefficients in italicized red are those that should reflect partial exposure to the program.

First, there are some patterns consistent with the dynamics of height-for-age over the life

course. Looking at children in municipalities that got Seguro Program after 2009, for instance, their

height-for-age z-scores tend to be higher for older children. This is consistent with the existing literature

which suggests that in developing and middle-income countries, height for age tends to decline until

weaning before increasing and then leveling off.

Looking at the coefficients that should reflect some impact of the program, generally most are

insignificant (and a few are actually negative). However, as discussed above, simply reading these

coefficients out of the table will not provide program effects. Instead, we will have to estimate some of

the difference-in-difference comparisons that were outlined in the previous section. This will help

disentangle selective roll-out, age effects, and time trends to see if there was in fact an effect of the

program. While there are a great number of possible comparisons that could be done, the few that

were discussed above are presented in Table 4. The first row in each panel of estimates includes all

children in the sample, while the second only includes children in families without anyone working in the

formal sector in MxFLS1. The third row in each panel includes children in families that had no one with

health insurance in MxFLS1. The fourth and fifth rows show results of quantile regressions which will be

discussed in the next section.

Some of the estimates of program effect (in Table 4) are positive and significant, but many

others are not different from zero (and some even have a negative sign). The largest estimate of

program effect is seen comparing children who got the program in 2002 to those who got the program

after 2009 (in the second column). The estimates do not change drastically when we limit the sample to

households most likely to benefit from the program in the second and third rows in each panel. The

estimates of full versus partial exposure are also all insignificant.

Estimating $(\alpha_{D1}^3 - \alpha_{A1}^3) - (\alpha_{D0}^3 - \alpha_{A0}^3)$ and $(\alpha_{D1}^4 - \alpha_{A1}^4) - (\alpha_{D0}^4 - \alpha_{A0}^4)$ can, as mentioned

above, serve as a robustness check on whether differential time trends are influencing the estimates in

the first column of panel A in Table 4, since there should be no effect of the program on these children.

However, these estimates are very small and insignificant, suggesting that differential time trends are

not affecting our estimates. Taken together, these difference-in-difference estimates provide some

evidence that the program did had an effect on child height, but it is primarily observed when we

compare children who got the program early to those who got it after 2008.

Quantile regressions

As discussed previously, the program might plausibly have had a greater impact on certain types

of households, which was the motivation for repeating the first regression but limiting it to the sample

20

of households in the informal sector or households without insurance at baseline. Here, we do something similar by performing a quantile regression (on the 25th and 75th percentiles) for the entire sample. The results of the difference-in-differences using this regression are presented in the fourth and fifth rows in each panel of Table 4. The results are not strikingly different than the results from the original ordinary least squares model. While we might have predicted that the health impact would be greater for children who are near the lower end of the height-for-age distribution, the impact is actually very similar for children at the 25th and 75th percentiles.

Other health outcomes

There are a number of other child and infant health outcomes and behaviors that are of interest that could potentially be affected by Seguro Popular. These effects are more difficult to estimate, however, because there is not the critical window for impact as there is for child height. There is some previous work on adults suggesting that Seguro Popular increased the use of primary care and health-seeking behavior but fairly little work looking specifically at use of care for children (Knox 2008, Gakidou et al 20011; Danese-DlSantos et all 20011, for instance). Future work could look at vaccination rates, neonatal care, or overall use of care to get at whether the program resulted in children receiving more care. However, without the benefit of a critical window as we have for child height, any results on these outcomes using observational data will be suggestive rather than causal. Given these caveats, very preliminary results on use of care suggests that overall health-seeking behavior declined substantially between 2002 and 2005 before increasing somewhat by 2009 in areas that got the program in 2002 (compared to those that got it after 2008). A similar but smaller decline is seen in the areas that got the program between 2003 and 2008. If this decline in use of care is related to the roll-out or implementation of the program, it provides some suggestion of a possible mechanism for why it took so long to see an effect of Seguro Popular on child height. However, considerable additional work is

needed to determine the cause of the fall in healthcare use and how it was related to the

implementation of the program.

Program change in 2006

At the start of 2006, a new set of health interventions and programs was added to Seguro

Popular that specifically focused on the health of infants and their mothers (as discussed in the

description of the program). One concern is that, in terms of its relevance for child health, Seguro

Popular might have been a fundamentally different program before and after the start of 2006. One

way to get at this issue would be to compare the effect of the program on child height before and after

the program change. Unfortunately, there is too little variation after 2006 due to the limited time

between 2006 and MxFLS3 (which began in 2009) to estimate the impact of the program with these

data.

Enrollment

The previous analyses were all versions of intent-to-treat, since there is likely selection into

enrollment. However, it is still of some interest to know what characteristics are associated with

enrollment to get suggestive evidence about which households actually enrolled in and benefited from

the program. Enrollment in Seguro Popular during MxFLS3 was used as an outcome in a linear

probability model (results are in Table 5). The unit of observation is now the household, since

households were enrolled in the program together. Households without children young children were

also included in these regressions. There are three versions of the model: one which includes basic

household information like per capita expenditure and health insurance coverage, a second which adds

household health characteristics, and a third which adds community health characteristics.

Generally, household resources (education and per capita expenditure in this model) are

negatively associated with being enrolled in the program. Having health insurance at baseline, as

expected, is also strongly and negatively related to the likelihood that households have Seguro Popular

at MxFLS3. The later the program was rolled out in their community, the less likely the household was to have enrolled, which is consistent with the results suggesting it took time for results of the program to develop. Most of the individual health characteristics are not significantly related to enrollment, other than the female household head having hypertension and the male household head having been diagnosed with cancer, both of which are positively associated with enrollment. Looking at the third column, households are more likely to be enrolled by 2009 if they live in a community with a SSA clinic (the clinics where Seguro Popular enrollees would receive care). They are also more likely to be enrolled if they live in a community with health care facilities providing key services like mother and child care. They generally are more likely to be enrolled if they live in communities with higher quality health care facilities, as measured by characteristics like having good ventilation and clean floors.

CONCLUSION AND DISCUSSION

The Mexican public health insurance program Seguro Popular was a large and ambitious attempt to ensure that all of Mexico's citizens had access to basic health care and services. It required a great deal of time and money to implement and resulted in significant investment in health care infrastructure. Many other middle-income countries have also implemented similar public health insurance programs, so lessons from Mexico have the potential to be relevant to other contexts. Despite this, very little is known about its effects, and particularly its impact on child health.   While there is some evidence that the program has had an effect on birth weight and infant mortality, these results are fairly sparse and often suffer from methodological shortcomings.

This analysis dealt with potential problems in this literature in a number of ways. The first was the decision to focus on child height as an indicator of child nutrition. Poor child nutrition has very real consequences both for health in childhood and for later-life outcomes like adult height, earnings, and longevity. It is also known from the nutrition and biology literature that height is affected by nutrition during a critical period in early childhood. Using this fact, in combination with the gradual roll-out of the

23

program and the panel nature of the data, allows the comparison of children who would have benefited from the program with children who would not have benefited in a modified difference-in-difference analysis. The critical window for child height allows us to estimate a causal effect using observational data, which would not be possible for other health outcomes that do not share this trait.

Another challenge with this type of analysis is that public programs like Seguro Popular are generally not introduced randomly, and are often supposed to reach areas with greater need first. While this was officially the case with Seguro Popular as well, in practice it is more complicated and not at all clear how roll-out was actually implemented. Combining communities into four groups based on when the program was first introduced helps to deal with this issue if communities that got the program around the same time are more likely to be similar in terms of the characteristics that determined the order of the program's introduction. Including community fixed-effects sweeps out any time-invariant community traits.

Another issue to be dealt with is that families who choose to enroll in a public health insurance program are likely different from families who do not, and plausibly in ways that are related to our health outcomes of interest. For this reason, all the estimates are intent-to-treat. Running the same regressions for the full sample of children and comparing them to the estimates for children whose families worked entirely in the informal sector at baseline and to those for children whose families did not have health insurance at baseline helps to get a sense of who might have benefited from the program and whether the estimates are being driven down by households who could not benefit from the program. However, our estimates are generally similar for the full sample and for the selected sub-samples.

The results provide little evidence for large or wide-spread benefits of the program in terms of child height. There are some positive effects of the program, but they are no stronger for the children we would initially expect to benefit the most. In fact, the quantiles regressions suggest they may be

concentrated among the upper end of the distribution of height-for-age rather than the bottom. In general, the results seem to be strongest for children who were exposed to the program their entire lives and who lived in areas that had had the program established for many years.

There are a number of possible explanations for these results. It could be, for instance, that it took time for the program to become fully established and for the infrastructure to catch up with the program requirements. This would be consistent with the preliminary, suggestive evidence on use of care which shows that health-seeking behavior initially fell in MxFLS2 communities that got the program earlier (compared to those communities that got the program after 2008). While some commentators have suggested anecdotally that there were shortages and long wait times, this has not been empirically demonstrated. Future work could try to get at this issue by analyzing use of care, wait times, costs per visit, and also community characteristics like number of clinics and doctors. It could also be that the program had sufficient infrastructure and employees but was better run and managed after it had been in place for some length of time. Alternatively, it may have taken some time for demand to increase as households learned about the program.

The existence of the Oportunidades program in Mexico might provide another possible explanation for a lack of clear effect of Seguro Popular. There is significant overlap in the populations covered by the two programs, and Oportunidades already provides many services that might affect the health of infants, young children, and their mothers. For these families, Seguro Popular may have more of an effect on healthcare spending and resources within the household in the event of an extreme health shock, while having less of an effect on health and health care for very young children. Unfortunately, these possible supply and demand explanations cannot be easily disentangled with these data.

These findings have serious implications given that evaluations of policies and randomized controlled trials are often done in the very short term. These results suggest that policy makers should

anticipate the possibility that it may take years for significant results to be seen or measured. This may be particularly true in contexts where implementation and management of new programs is imperfect or infrastructure needs to be scaled up.

The composition of the households who enrolled at each point in time may also have changed, causing some of the patterns that are observed if, for instance, the children who might have benefited the most were reached later than better off children. It may be that giving states explicit enrollment goals encouraged municipalities to focus on enrolling the easy-to-reach households first. It is also possible that it took time for household to learn about the program and to enroll, and households with more resources and higher levels of education were able to do this more quickly. Whatever the underlying reason, this pattern of reaching better off children first might be one factor limiting the program's initial effectiveness.

Overall, this analysis provides evidence of a limited causal impact of the program on child height. It also suggests that the impact took several years to be established. Although some preliminary evidence on possible explanations for this small effect is provided, more research is needed to definitively disentangle the possible mechanisms. This is particularly important if other countries look to Mexico's experience for lessons on public health provision. All of this points to a need for researchers and policy makers to think harder about how similar programs are implemented and run, how beneficiaries are targeted, and how long program evaluators might have to wait to see any real health benefits.

REFERENCES

Atun, Rifat, Luiz Odorico Monteiro de Andrade, Gisele Almeida, Daniel Cotlear, T Dmytraczenko, Patricia Frenz, Patricia Garcia, Octavio Gómez-Dantés, Felicia Knaul, Carles Muntaner, Juliana Braga de Paula, Felix Rigoli, Pastor Castell-Florit Serrate, and Adam Wagstaff (2015). "Health-System Reform and Universal Health Coverage in Latin America." The Lancet 385: 1230-47.

Aquilera, Nelly, and Grecia Marrufo (2006). "New Evidence of the Effects of Publicly Provided Insurance on Health Outcomes: The Case of the Seguro Popular." Presentation at LACEA.

Azuara, Oliver, and Ioana Marinescu (2010). "Informality and the Expansion of Social Protection Programs. Evidence from Mexico." Journal of Health Economics 32(5): 938-950.

Avila-Burgos, Leticia, Edson Servan-Mori, Veronica Wirtz, Sandra Sosa-Rubi, and Aaron Salinas-Rodriguez (2013) "Effect of Seguro Popular on Health Expenditures in Mexican Households Ten Years after Its Implementation." Salud Publica de Mexico 55(2): S91-S99.

Barros, Rodrigo (2008). "Wealthier but Not Much Healthier: Effects of a Health Insurance Program for the Poor in Mexico." Working paper.

Bleich, Sara N., David M. Cutler, Alyce S. Adams, Rafael Lozano, and Christopher J. L. Murray (2007). "Impact of Insurance and Supply of Health Professionals on Coverage of Treatment for Hypertension in Mexico: Population Based Study." British Medical Journal 335(7625): 875-B878.

Bosch, Mariano and Raymundo M. Campos-Vázquez (2010). "The Trade-Offs of Social Assistance Programs in the Labor Market: The Case of the 'Seguro Popular' Program in Mexico." Working paper.

Bosch, Mariano, M. Belen Cobacho, and Carmen Pages (2012). "Taking Stock of Nine Years of Implementation of Seguro Popular in Mexico." The Inter-American Development Bank.

Campos-Vázquez, Raymundo M. and Melissa A. Knox (2011). "Social Protection Programs and Employment: The Case of Mexico's Seguro Popular Program." Economia Mexicana – Nueva Epoca 22(2).

Comisión Nacional de Protección Social en Salud (2012). "Catálogo Universal de Servicios de Salud CAUSES 2012). Gobierno Federal <www.gobiernofederal.gob.mx>.

Chavez-Martin del Campo, Juan Carlos (2006). "Does Conditionality Generate Heterogeneity and Regressivity in Program Impacts? The Progresa Experience." Working paper.

Conti, Gabriella and Rita Ginja (2014). "Evaluating a Universal Health Insurance Program: Evidence from Mexico." Working paper.

Currie, Janet and Jonathan Gruber (1996a). "Saving Babies: The Efficacy and Cost of Recent Changes in the Medicaid Eligibility of Pregnant Women." Journal of Political Economy 104(6): 1263-1296.

Currie, Janet and Jonathan Gruber (1996b). "Health Insurance Eligibility, Utilization of Medical Care, and Child Health." The Quarterly Journal of Economics 111(2): 431-466.

Currie, Janet and Tom Vogl (2013). "Early-Life Health and Adult Circumstance in Developing Countries." Annual Review of Economics 5:1-36.

Dammert, Ana (2008). "Heterogeneous Impacts of Conditional Cash Transfers: Evidence from Nicaragua." IZA Discussion Paper No. 3653.

Danese-dlSantos, Laure, Sandra Sosa-Rubi, and Atanacio Valencia-Mendoza (2011). "Analysis of Changes in the Association of Income and the Utilization of Curative Health Services in Mexico between 2000 and 2006." BMC Public Health 11: 771.

Djebbari, Habiba and Jeffrey Smith (2008). "Heterogeneous Impacts in PROGRESA." Journal of Econometrics 145: 64-80.

Duflo, Esther (1999). "Grandmothers and Granddaughters: Old-Age Pensions and Intrahousehold Allocation in South Africa." World Bank Economic Review 17(1): 1-25.

Duflo, Esther (2000). "Child Health and Household Resources in South Africa: Evidence from the Old Age Pension Program." American Economic Review 90(2): 393-398.

Duflo, Esther (2001). "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment." The American Economic Review 91(4): 795-813.

Hernández, Robert Duval and Ricardo Smith Ramírez (2011). "Informality and Seguro Popular under Segmented Labor Markets." Working paper.

Farfan, Gabriela, Maria Genoni, Luis Rubalcava, Graciela Teruel, and Duncan Thomas (2013). "Oportunidades and its Impact on Child Nutrition." Working paper.

Frankenberg, Elizabeth, Wayan Suriastini, and Duncan Thomas (2005). "Can Expanding Access to Basic Healthcare Improve Children's Health Status? Lessons from Indonesia's 'Midwife in the Village' Programme." Population Studies 59(1): 5-19.

Frenk, Julio (2006). "Bridging the Divide: Global Lessons from Evidence-Based Healthy Policy in Mexico." <u>Lancet</u> 368: 954-961.

Frenk, Julio, Eduardo González-Pier, Octavio Gómez-Dantés, Miguel A. Lezana, Felicia Marie Knaul (2006). "Comprehensive Reform to Improve Health System Performance in Mexico." <u>Lancet</u> 368: 1524-34.

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

Galárraga, Omar, Sandra G. Sosa-Rubí, Aarón Salinas-Rodriguez, and Sergio Sesma-Vázquez (2010). "Health Insurance for the Poor: Impact on Catastrophic and Out-of-Pocket Expenditures in Mexico." <u>European Journal of Health</u> 11(5): 437-447.

Garcia-Diaz, Rocio and Sandra G. Sosa-Rubí (2011). "Analysis of the Distributional Impact of Out-of-Pocket Health Payments: Evidence from a Public Health Insurance Program for the Poor in Mexico." <u>Journal of Health Economics</u> 30: 707-718.

Haddad, Lawrence J. and Bouis Howarth E. (1991). "The Impact of Nutritional Status on Agricultural Productivity: Wage Evidence from the Phillippines." <u>Oxford Bulletin of Economics and Statistics</u> 53: 45-68.

Hernández, Robert Duval (2011). "Informality and Seguro Popular under Segmented Labor Markets." Working paper.

Imai, Kosuke, Gary King, and Clayton Nall (2009). "The Essential Role of Pair Matching in Cluster-Randomized Experiments, with Application to the Mexican Universal Health Insurance Evaluation." <u>Statistical Science</u> 24(1): 29-53.

Kaestner, Robert, Theodore Joyce, and Andrew Racine (1999). "Does Publicly Provided Health Insurance Improve the Health of Low-Income Children in the United States?" NBER working paper.

King, Gary, Emmanuela Gakidou, Nimala Ravishankar, Ryan T. Moore, Jason Lakin, Manett Vargas, Martha María Téllez-Rojo, Juan Eugenio Hernández Ávila, Mauricio Hérnández Ávila and Héctor Hernández Llamas (2007). "A Politically Robust' Experimental Design for Public Policy Evaluation, with Application to the Mexican Universal Health Insurance Program." <u>Journal of Policy Analysis and Management</u> 26(3): 479-506.

King, Gary, Emmanuela Gakidou, Kosuk e Imai, Jason Lakin, Ryan T. Moore, Clayton Nall, Nirmala Ravishankar, Manett Vargas, Martha María Téllez-Rojo, Juan Eugenio Hernández Ávila, Mauricio Hernández Ávila, Héctor Hernández Llamas (2009). "Public

Policy for the Poor? A Randomised Assessment of the Mexican Universal Health Insurance Programme." Lancet 373: 1447-1454.

Knaul, Felicia Marie, Héctor Arreola-Ornelas, Oscar Méndez-Carniado, Chloe Bryson-Cahn, Jeremy Barofsky, Rachel Macguire, Martha Miranda, and Sergio Sesma (2006). "Evidence is Good for Your Health System: Policy Reform to Remedy Catastrophic and Impoverishing Health Spending in Mexico." Lancet 368: 1828-1841.

Knaul, Felicia Marie, Eduardo González-Pier, Octavio Gómez-Dantés, David García-Junco, Héctor Arreola-Ornelas, Mariana Barraza-Lloréns, Rosa Sandoval, Francisco Caballero, Mauricio Hernández-Avilla, Mercedes Juan, David Kershenobich, Gustavo Nigenda, Enrique Ruelas, Jaime Sepúlveda, Roberto Tapia, Guillermo Soberón, Salomón Chertorivski, and Julio Frenk (2012). "The Quest for Universal Health Coverage: Achieving Social Protection for All in Mexico." The Lancet 380(9849): 1259-1279.

Knox, Melissa (2008). "Health Insurance for All: An Evaluation of Mexico's Seguro Popular Program." Working paper.

Lee, Soohyung and Azeem Shaikh (2014). "Multiple Testing and Heterogeneous Treatment Effects: Re-Evaluating the Effect of Progresa on School Enrollment." Journal of Applied Econometrics 29: 612-626.

Leyva-Flores, Rene, Edson Servan-Mori, Cesar Infante-Xibille, Blanca Estela Pelcastre-Villafuerte, and Tonatiuh Gonzalez (2014). "Primary Health Care Utilization by the Mexican Indigenous Population: The Role of the Seguro Popular in Socially Inequitable Contexts." PLOS One 9(8).

Martorell, Reynaldo and Jean-Pierre Habicht (1986). "Growth in Early Childhood in Developing Countries." In Frank Falkner and J.M. Tanner (eds.) *Human Growth: A Comprehensive Treatise*, Vol. 3, New York: Plenum Press, pp. 241-262.

Muñoz-Hernández, Onofre, Salomón Chertorivski-Woldenberg, Gabriel Cortés-Gallo, and Ricardo Pérez-Cuevas (2012). "The Medical Insurance for a New Generation: A Viable Answer for the Health Needs of Mexican Children." Salud Pública de México 54(sup 1): 3-10.

Newhouse, Joseph P. and the Insurance Experiment Group (2003). "Free for All? Lessons from the RAND Health Insurance Experiment." *Harvard University Press*.

National Commission of Social Health Protection (2012-2013). "Medican Insurance for a New Generation/Insurance 21st Century." <http://portal.salud.gob.mx/codigos/columnas/evaluacion_programas/fichas/12_S201_FM.pdf>

Pfutze, Tobias (2014). "The Effects of Mexico's Seguro Popular Health Insurance on Infant Mortality: An Estimation with Selection on the Outcome Variable." World Development 59: 475-486.

Rubalcava, Luis and Graciela Teruel (2006). "Mexican Family Life Survey, First Wave." Working Paper, www.ennvih-mxfls.org.

Rubalcava, Luis and Graciela Teruel (2008). "Mexican Family Life Survey, Second Wave." Working Paper, www.ennvih-mxfls.org.

Rubalcava, Luis and Graciela Teruel (2013). "Mexican Family Life Survey, Third Wave." Working Paper, www.ennvih-mxfls.org.

Rodríguez-Ortega, Evelyne Eugenia and Elviro Mireya Pasillas-Torres (2012). "Family Healthcare Expenditure on Children Affiliated to the Medical Insurance for a New Generation." Salud Pública de México 54(sup 1): 65-72.

Scott, John (2006). "Seguro Popular Incidence Analysis." In *Chap 3. pages 147-166 of: Decentralized Service Delivery for the Poor,* vol. II. World Bank Publications.

Sosa-Rubí, Sandra G., Omar Galárraga, and Jeffrey E. Harris (2008). "Heterogeneous Impact of the 'Seguro Popular' Program on the Utilization of Obstetrical Services in Mexico, 2001-2006: A Multinomial Probit Model with a Discrete Endogenous Variable." Journal of Health Economics 28:20-34.

Sosa-Rubí, Sandra G., Omar Galárraga and Ruy López-Ridaura (2009). "Diabetes Treatment and Control: The Effect of Public Health Insurance for the Poor in Mexico." Bulletin of the World Health Organization 87:512-519.

Sosa-Rubí, Sandra G., Aaron Salinas-Rodriguez, and Omar Galarraga (2011). "Impact of 'Seguro Popular' on Catastrophic and Out-of-Pocket health Expenditures in Rural and Urban Mexico, 2005-2008." Salud Publica de Mexico 53(4): S425-S435.

Spenkuch, Jorg L (2012). "Moral Hazard and Selection among the Poor: Evidence from a Randomized Experiment." Journal of Health Economics 31: 72-85.

Steckel, Richard H. (1995). "Stature and the Standard of Living." Journal of Economic Literature 33: 1903-1940.

Strauss, John and Duncan Thomas (1998). "Health, Nutrition and Economic Development." Journal of Economic Literature 36: 766-817.

Strauss, John and Duncan Thomas (1995). "Human Resources: Empirical Models of Household

Decisions." Pp. 1885-2023 in *Handbook of Development Economics*, Vol. IIIA, edited by J.R. Behrman and T.N. Srinivasan. Amsterdam: North Holland.

Wirtz, Veronika J., yared Santa-Ana-Tellez, Edson Servan-Mori, and Leticia Avila-Burgos (2012). "Heterogeneous Effects of Health Insurance on Out-of-Pocket Expenditure on Medicines in Mexico." <u>Value in Health</u> 15: 593-603.

TABLES AND FIGURES

Table 1: Community characteristics in 2002 by date of roll-out	
Panel A – Average education of household head	
Received SP by 2002	6.3
Received SP by 2005	6.2
Received SP by 2008	4.9
Received SP after 2008	8.1
Panel B – Average log of per capita expenditure	
Received SP by 2002	6.9
Received SP by 2005	6.8
Received SP by 2008	6.5
Received SP after 2008	7.1
Panel C – Percent of adults informally employed	
Received SP by 2002	60%
Received SP by 2005	63%
Received SP by 2008	69%
Received SP after 2008	60%
Panel D – Percent of adults without insurance	
Received SP by 2002	45%
Received SP by 2005	54%
Received SP by 2008	68%
Received SP after 2008	53%

Table 2: Descriptive statistics of the sample								
	MxFLS 1 (2002)		MxFLS 2 (2005)		MxFLS 3 (2009)		Full sample	
	Age: 0-2	Age 3-9	Age 0-2	Age 3-9	Age 0-2	Age 3-9	Age 0-2	Age 3-9
Panel A – Average height-for-age z-scores								
Received SP by 2002	-0.48	-0.28	-0.40	-0.31	0.11	-0.18	-0.16	-0.26
Received SP by 2005	-0.66	-0.59	-0.48	-0.45	-0.37	-0.33	-0.47	-0.45
Received SP by 2008	-0.75	-0.99	-0.61	-0.80	-0.55	-0.72	-0.62	-0.83
Received SP after 2008	-0.50	-0.50	-0.45	-0.30	-0.20	-0.53	-0.46	-0.43
Overall	-0.61	-0.58	-0.49	-0.47	-0.30	-0.39	-0.44	-0.49
Panel B – Household and child characteristics								
Age in years	4.73		4.66		4.32		4.57	
Education of household head	6.31		6.36		6.66		6.49	
Log per capita expenditures	6.61		6.31		7.05		6.71	
Household size	5.74		5.80		5.87		5.87	

Table 3: Height-for-age z-scores (Full sample of children 0 to 9 years old at each wave)			(1)	(2)	(3)	(4)
			Received SP in 2002	Received SP in 2003-2005	Received SP in 2006-2008	Not received SP by 2009
Notation	Period	Age group				
α_{A0}^{v}	2002	A: 7 to 9	0.04	-0.05	-0.22**	-0.26***
			(0.08)	(0.07)	(0.10)	(0.10)
α_{B0}^{v}		B: 5 to 6	-0.07	0.02	-0.17	-0.11
			(0.09)	(0.08)	(0.12)	(0.09)
α_{C0}^{v}		C: 3 to 4	-0.12	-0.07	-0.22*	-0.34***
			(0.11)	(0.08)	(0.11)	(0.09)
α_{D0}^{v}		D: 0 to 2	-0.21*	-0.12	0.07	-0.90***
			(0.11)	(0.11)	(0.13)	(0.11)
α_{A1}^{v}	2005	A: 7 to 9	-0.20	-0.12	-0.27**	0.51***
			(0.11)	(0.09)	(0.11)	(0.11)
α_{B1}^{v}		B: 5 to 6	*-0.22*	*-0.15*	-0.21	0.22**
			(0.15)	*(0.11)*	(0.14)	(0.31)
α_{C1}^{v}		C: 3 to 4	*-0.15*	*0.08*	0.06	-0.31***
			(0.14)	*(0.11)*	(0.14)	(0.11)
α_{D1}^{v}		D: 0 to 2	**-0.14**	*0.03*	0.12	-0.89***
			(0.19)	*(0.12)*	(0.16)	(0.13)
α_{A2}^{v}	2009	A: 7 to 9	*0.26*	*0.29*	*0.28*	0.14
			(0.21)	*(0.21)*	*(0.22)*	(0.22)
α_{B2}^{v}		B: 5 to 6	**-0.51****	*-0.26*	*-0.51**	0.43*
			(0.26)	*(0.25)*	*(0.26)*	(0.25)
α_{C2}^{v}		C: 3 to 4	**0.16**	**0.28**	*0.13*	-0.46**
			(0.19)	**(0.19)**	*(0.20)*	(0.19)
α_{D2}^{v}		D: 0 to 2	**0.09**	**-0.07**	*-0.06*	-0.68***
			(0.25)	**(0.25)**	*(0.26)*	(0.26)
		F-statistic [column] (p-value)	1.53 (0.11)	1.04 (0.41)	2.29 (0.01)	11.09 (0.00)
		F-statistic [full model] (p-value)	8.58 (0.00)			
		Observations	17,776			

This table is the result of running one ordinary least squares regression on the sample of children age 0 to 9 at each wave of MxFLS. The regression also includes gender, age in months, and municipality fixed effects. Robust standard errors are clustered at the municipality-year-cohort level (clustering at the municipality, municipality-time, or municipality-cohort levels give essentially the same result). Numbers in blue reflect the full impact of the program, while numbers in red reflect partial exposure.

Table 4: Difference-in-difference estimates of the program impact		
Panel A: Full program effect		
	(1) $(\alpha_{D1}^{1} - \alpha_{A1}^{1}) - (\alpha_{D0}^{1} - \alpha_{A0}^{1})$	(2) $(\alpha_{D1}^{1} - \alpha_{A1}^{1}) - (\alpha_{D1}^{4} - \alpha_{A1}^{4})$
Full sample	0.31	1.43***
	(0.24)	(0.32)
Informal sample	0.30	1.58***
	(0.39)	(0.53)
No insurance sample	0.15	1.06**
	(0.36)	(0.48)
Quantile regression (25%)	0.55	1.18***
	(0.32)	(0.37)
Quantile regression (75%)	0.30	1.38***
	(0.32)	(0.46)
Panel B: Full versus partial effect		
	(1) $(\alpha_{D2}^{1} - \alpha_{A2}^{1}) - (\alpha_{D2}^{4} - \alpha_{A2}^{4})$	(2) $(\alpha_{D2}^{2} - \alpha_{A2}^{2}) - (\alpha_{D2}^{4} - \alpha_{A2}^{4})$
Full sample	0.65	0.47
	(0.63)	(0.62)
Informal sample	-0.41	-0.81
	(0.70)	(0.69)
No insurance sample	1.59	1.46
	(1.18)	(1.16)
Quantile regression (25%)	-0.41	-0.51
	(0.51)	(0.51)
Quantile regression (75%)	1.05	0.98
	(0.67)	(0.66)

To get the program impact estimates in each cell, first the ordinary least squares regression outlined in the empirical specification was calculated to obtain the coefficients, which were then used to construct the difference-in-difference estimates. Nine versions of this model were run (each row in each panel represents one version). Robust standard errors are clustered at the municipality-year-cohort level (clustering at the municipality, municipality-time, or municipality-cohort levels give essentially the same result).

Table 5 – Household was enrolled in Seguro Popular by 2009 (100 = yes)

	(1) Basic	(2) Basic + health	(3) Basic + health + community
Basic household characteristics			
Education of the male head	-1.09*** (0.16)	-1.06*** (0.16)	-0.71*** (0.15)
Education of the female head	-0.50*** (0.18)	-0.48*** (0.18)	-0.28* (0.16)
Age of the male head	-0.11 (0.06)	-0.11 (0.06)	-0.09 (0.07)
Age of the female head	-0.04 (0.07)	-0.05 (0.07)	-0.08 (0.07)
No one in the household works in the formal sector	0.11 (0.91)	0.17 (0.94)	-1.23 (1.07)
The household has health insurance	-18.45*** (1.46)	-18.54*** (1.47)	-15.19*** (1.33)
Seguro Popular arrived in 2003 (omitted category is 2002)	-3.95 (4.80)	-4.00 (4.78)	-3.49 (5.70)
Seguro Popular arrived in 2004	0.82 (4.09)	0.72 (4.11)	0.46 (4.82)
Seguro Popular arrived in 2005	-6.58 (4.74)	-6.65 (4.78)	-5.94* (3.46)
Seguro Popular arrived in 2006	0.56 (5.78)	0.56 (5.75)	3.85 (5.53)
Seguro Popular arrived in 2007	-14.29*** (7.18)	-14.65** (7.13)	-9.43 (7.82)
Seguro Popular arrived in 2008	-38.81*** (6.94)	-38.81*** (6.88)	-31.94*** (10.26)
Seguro Popular arrived after 2008	-24.13*** (3.40)	-24.11*** (3.42)	-19.69*** (3.00)
Log of per capita expenditure	-7.48*** (0.99)	-7.29*** (0.97)	-4.53*** (0.94)
Household health characteristics			
Height of the male head		0.06 (0.08)	0.07 (0.08)
Height of the female head		-0.03 (0.07)	-0.04 (0.08)
If the male head smokes		-2.06 (1.76)	-0.32 (1.88)
If the female head smokes		1.76 (1.32)	1.44 (1.45)
If the male head exercises		-0.46 (1.73)	-1.01 (1.68)
If the female head exercises		-1.71 (1.21)	-1.08 (1.26)
If the male head has been diagnosed with hypertension		-0.69 (2.31)	1.19 (2.30)
If the female head has been diagnosed with hypertension		3.78** (1.52)	2.77* (1.63)
If the male head has been diagnosed with cancer		21.38* (12.00)	15.53 (14.90)
If the female head has been diagnosed with cancer		1.95 (3.73)	0.75 (3.80)
Community healthcare characteristics			
At least one SSA institution			3.35 (3.26)
At least one IMSS/ISSTE institution			-7.10 (4.50)
At least one private institution			-7.16** (3.54)
At least one other type of institution			-6.87** (2.70)
At least one center with a family planning area			3.43 (3.84)
At least one center with mother-child services area			4.06 (5.00)
At least one center with good ventilation in the exam room			10.57* (5.30)
At least one center with a very clean floor in the exam room			3.54 (3.54)
At least one center with a lab			-6.09 (4.86)
At least one center without frequent blackouts			0.54 (2.31)
At least one center with community outreach			0.45 (6.12)
Constant	104.56*** (10.41)	113.54*** (11.96)	79.46*** (17.10)
F-statistic (p-value)	35.63 (0.00)	68.92 (0.00)	871.97 (0.00)
Adjusted R-squared	0.22	0.23	0.24
Number of observations	7,957	7,957	6,119

Additional covariates in both models not shown in the table include whether the female and male head exist; number of household members of each gender from 0-4 years old, 5-9 years old, 10-14 years old, 15-24 years old, 25 to 64 years old, and 65 or older; whether someone in the household works; and state dummies. Additional covariates in model (2) not shown in the table include waist circumference, BMI, and blood pressure of the household heads and whether the household heads have been diagnosed with diabetes or heart disease. All dependent variables are measured at baseline. Robust standard errors clustered by municipality are in parenthesis.

Figure 1: Percent of children aged 0-2 with a cough in previous 4 weeks
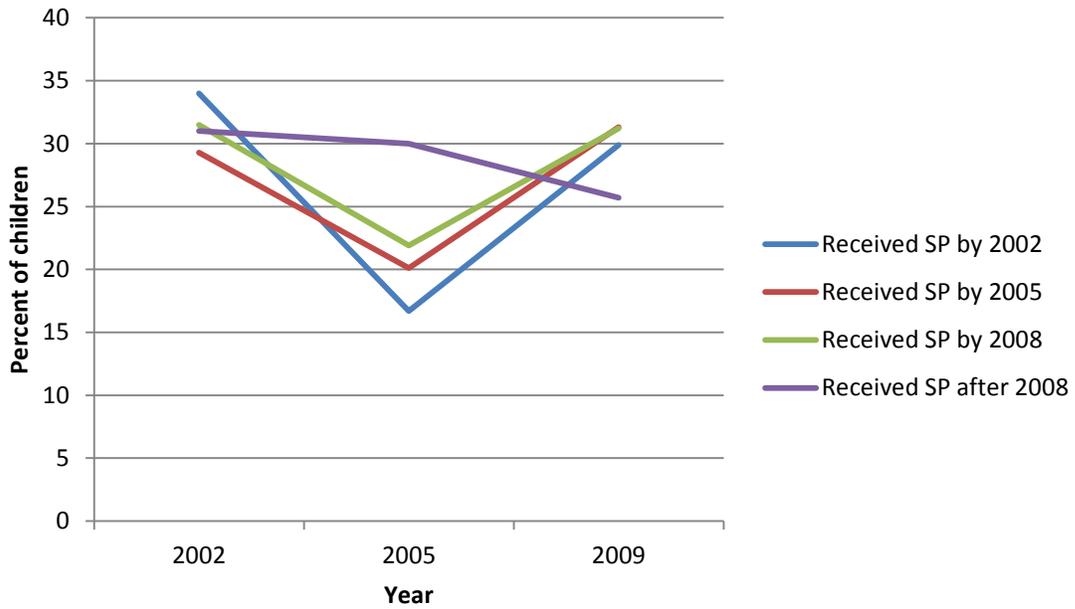
Figure 2: Percent of children aged 0-2 with diarrhea in previous 4 weeks
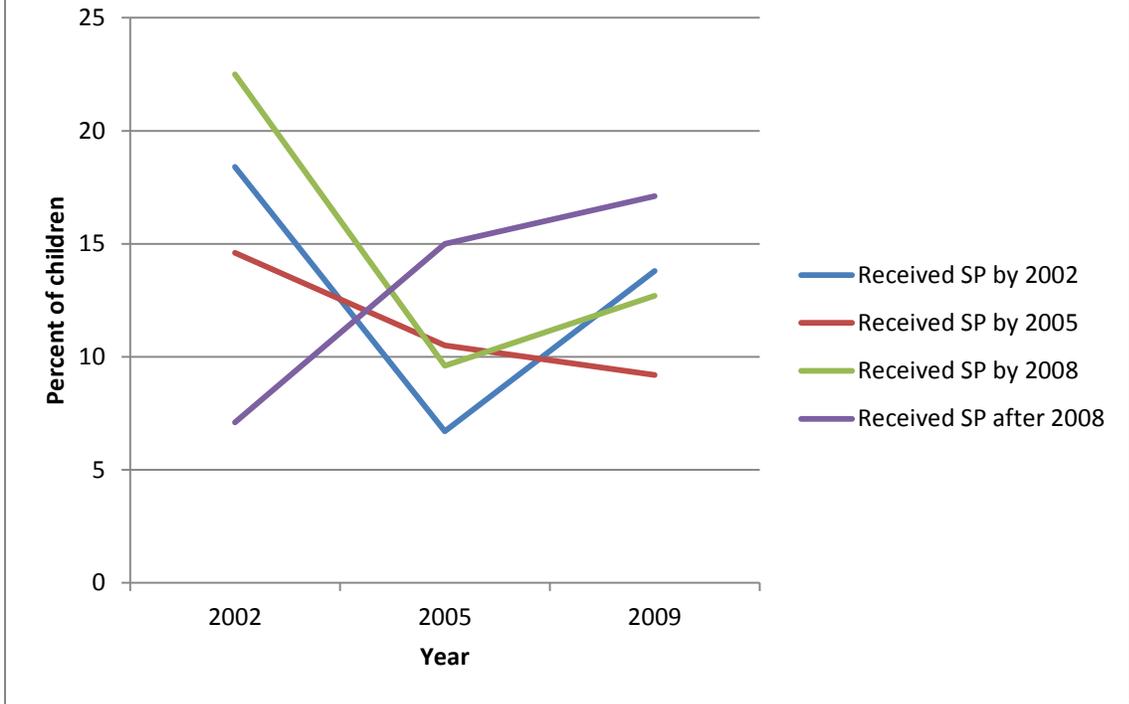
Figure 3: Expected impact of Seguro Popular on height-for-age z-scores				
	(1) Received SP in 2002	(2) Received SP in 2003-2005	(3) Received SP in 2006-2008	(4) Not received SP by 2009
2002: 7 to 9 years old	None	None	None	None
2002: 5 to 6 years old	None	None	None	None
2002: 3 to 4 years old	None	None	None	None
2002: 0 to 2 years old	None	None	None	None
2005: 7 to 9 years old	None	None	None	None
2005: 5 to 6 years old	Some	Some	None	None
2005: 3 to 4 years old	Some	Some	None	None
2005: 0 to 2 years old	Full	Some	None	None
2009: 7 to 9 years old	Some	Some	Some	None
2009: 5 to 6 years old	Full	Some	Some	None
2009: 3 to 4 years old	Full	Full	Some	None
2009: 0 to 2 years old	Full	Full	Some	None