The Effects of Mexicos Seguro Popular Health Insurance on Infant Mortality: An Estimation with Selection on the Outcome Variable. *

Tobias Pfutze

Oberlin College & American University

tpfutze@oberlin.edu

10 N. Professor St., Rice Hall 233, Oberlin, OH 44074, USA

*I would like to thank Natalia Radchenko and Amos Golan for useful advice. Computing resources used for this work provided by the American University High Performance Computing System; see www.american.edu/hpc for information on the system and its uses.

Abstract

This paper estimates the effects of Mexico's non-contributory health insurance on infant mortality during its first five years of roll-out, using data on births reported in the micro sample from Mexico's 2010 general population census. However, births of surviving children are more likely to be observed than births of non-surviving ones. This selection on the outcome variable is addressed using the weighted exogenous sampling maximum likelihood (WESML) estimator, developed by Manski and Lerman (1977) for the case of choice-based samples. The results indicate that the program can be expected to reduce Mexico's infant mortality by close to 5 out of 1,000 births.

JEL Classification: C81, I13, I15, I18
Keywords: Infant Mortality, Health Insurance, Mexico, Sample Selection

1 Introduction

Over the course of the last decade, many middle-income countries have implemented new policies aimed at easing social inequalities through the provision of basic government services to the population at the bottom of the income distribution. Often considered to have originated in cash transfer programs, such as Bolsa Escola in Brazil or Progresa in Mexico, these policies have been extended more recently to cover areas such as pensions and health care.

The present paper analyzes the effect of one such programs, the Seguro Popular health insurance in Mexico, on one of the most important public health outcomes, infant mortality. Exploiting variations in the program's lengthly roll-out, it will be shown that it had a significant impact on reducing neonatal and infant mortality. Moreover, coverage during pregnancy is a more important factor than coverage at birth. The analysis uses individual level survey data, which, as is often the case, suffers from potentially severe bias due to selection on the outcome variable. Not only is complete data for non-surviving children more likely to be missing, but data is only available on a mother's last live birth. Both aspects result in a higher probability that the birth of a surviving child forms part of the sample than that of a non-surviving one. This problem is addressed by using a simple correction method first proposed by Manski and Lerman (1977) for the case of choice-based sampling.

Selection on the dependent variable is a fairly common, but often ignored, problem when working with survey based mortality data. To my knowledge, this is the first paper that applies choice-based sampling techniques to the estimation of mortality risk to address this issue. It is also the first paper to rigorously estimate the impact of Seguro Popular on measures of child mortality. A few studies, however, have looked at the program's impact on the availability, and use of, particular interventions known to reduce mortality risks. A commissioned process evaluation (Muñoz Hernández, Cortes-Gallo, Pérez-Cuevas, Jasso-Gutiérrez, and Walsh 2010) focuses on the available care for premature births at providers accredited by Seguro Popular, and also provides some general estimates on neonatal mortality in Mexico. Sosa-Rubi, Galarraga, and Harris (2009) find that the program increased the proportions of babies delivered at Seguro Popular accredited clinics, reducing such deliveries at non-accredited private and public clinics.

Most of the emerging literature on non-contributory social protection programs has focused on a number of Latin American countries, and especially on Mexico. This focus is mostly the result of the large number of programs in the region, combined with comparatively good and easily accessible data. An early and very influential discussion on such programs, and in particular of Seguro Popular, can be found in Levy (2008). His main concern is that the emergence of social protection programs aimed at households in the informal sector will decrease the incentives for beneficiaries to find formal sector employment (where they would have to pay payroll taxes to enjoy similar services) and, as a result, act as a break on economic development. Consequently, a large share of the empirical literature of interest has focused on this particular question. For the case of Seguro Popular, Barros (2008) does not find any statistically significant effect of the program on informality. Azuara and Marinescu (2010) find that it increased informality only for a few subgroups of the population and by less than 2%. Overall, they do not find any effect on the transition probabilities between formal and informal occupations. Similarly, Martínez and Aguilera (2010) find a statistically significantly negative effect on formality, but of small magnitude. However, Bosch and Campos-Vázquez (2010) find that the program significantly reduced the creation of formal employment in small and medium sized firms. For the case of Colombia, Camacho, Conover, and Hoyos (2010) find that the expansion of government provided health care increased informality by 2-5%.

In a paper closely related to this one, Camacho and Conover (2013) estimate the effect of the subsidized Colombian health insurance on health outcomes of newborns in a single metropolitan area using a regression discontinuity around the eligibility threshold. Their estimates show that the program significantly reduced the incidence of a low birth weight and increased the Apgar score. Concerning other outcomes, Barros (2008) finds that Seguro Popular decreased beneficiaries out of pocket expenses on health services, but that it did not improve health outcomes. Similarly, Grogger, Arnold, León, Ome, and Triyana (2010) also find that the program did reduce household expenditures on health, but differently in urban and rural areas. Whereas in the former the reduction is mostly due to savings on prescription drugs, in the latter the results are driven by fewer expenses on catastrophic events. A lack of improvements in health outcomes is also the result in Duval-Hernández and Smith-Ramírez (2011), who conclude that the program's costs are therefore not justified. These results are also discussed in Scott and Aguilera (2010), who point out that at this point it is still impossible to estimate the program's effect on long-term health outcomes. In this context, the findings by Ruvalcaba and Parker (2010) are of particular interest. Using self-reported health data and matching methods, it is found that while Seguro Popular does not reduce the incidence of chronic diseases such as diabetes (which are unlikely to be affected in the short run), it did have a significant effect on the reduction in cholesterol, and, in some case, high blood pressure. Given the long-term nature of many health outcomes in the adult population, a focus on results that can be expected to manifest themselves in the short term, such as infant mortality, seems to be a fruitful exercise.

Over the course of the last two decades, Mexico's mortality rates for children under five years of age have been on a slow decline. Using UNICEF data

[1], figure 1 shows the evolution of child, infant and neonatal mortality since 1990. While it would be interesting to include child mortality (i.e. mortality during the first five years of life) into the present study, data limitations make a convincing analysis only feasible for infant and neonatal mortality. But given that most child mortality occurs shortly after birth, this should not be a practically important limitation. Based on the UNICEF estimates, over the course of the last twenty years, infant mortality in Mexico has decreased from 38.1 per one thousand live births in 1990 to 13.4 in 2011, the corresponding figures for neonatal mortality are 17.2 and 7.0. These numbers, however, mask a high degree of heterogeneity according to households' socioeconomic status. For example, Muñoz Hernández, Cortes-Gallo, Pérez-Cuevas, Jasso-Gutiérrez, and Walsh (2010) point out that for mothers affiliated to Mexico's public health insurance IMSS (see below), who are in their majority by no means prosperous, neonatal mortality declined between 1990 and 2005 from 12.47 to 8.7. For the population as a whole the numbers are 17.2 and 9.4 according to UNICEF. These figures suggest significantly higher rates for the uncovered population, but also some convergence between the two groups in recent years. It is a standard result in the public health literature that pre-natal health care is a strong determinant for child mortality risks, often mediated through a low birth weight (see, for example Gortmaker (1979) or Shi, Macinko, Starfield, Xu, Regan, Politzer, and Wulu (2004) among a bottomless literature). The present study confirms this common result.

[1] The data used can be found at www.childmortality.org

(Figure 1 about here)

This paper has five sections. The next one explains in greater detail the Seguro Popular program and puts it into the Mexican context. Section 3 discusses the data and empirical strategy. It also provides a detailed discussion on the choice-based sampling estimator used. Section 4 presents the result, including robustness checks, and section 5 concludes.

2 Description of the Program

Mexico essentially has a three tier health care system. At the top, most of the country's upper middle class enjoys the benefits of private insurance, often provided by employers in addition to the mandatory contributions to the public system. Next, the traditional mainstay of public health care system consists of various agencies catering to formally employed workers; the most important of which is the *Instituto Mexicano del Seguro Social* (IMSS). All formally employed private sector employees have to be affiliated to the IMSS, while employees in the public sector are covered by different bodies, depending on whether they work at the federal or state level. In addition, the armed forces and the national parastatal oil monopoly (PEMEX) also have their own agencies. At the bottom are most informal workers who traditionally could not count on any reliable public health services.

When the IMSS was founded in 1943, it was assumed that continued economic growth would eventually bring the entire labor force into the system. The provision of health care services for the non-covered population was therefore considered to be a transitory problem. Over the following decades, this population was served by a variety of different programs and clinics run directly by the Mexican Department of Health (*Secretaría de Salud*) or its state level equivalents. However, by the mid 2000s the IMSS covered only around one third of the country's population, with a range of around 50% in Mexico City and Chihuahua to less than 10% in the poor southern states of Oaxaca, Guerrero and Chiapas (Scott and Aguilera 2010).

By the early 2000s, after several decades without any progress of extending the coverage of IMSS, and given that half the country's population did not have access to any form of health insurance, the Mexican government decided to implement a universally accessible health insurance program aimed at the so-far uncovered population. The Seguro Popular started with a pilot program in 2001 and initiated its universal roll-out in 2004. The program is jointly implemented by the federal and state governments, and its costs are shared between the two. For each individual affiliated, a fixed quota (around 12% of a yearly minimum wage) is paid into the system, of which the federal government pays $\frac{5}{6}$, and the state government $\frac{1}{6}$. In addition, beneficiaries may have to contribute based on an evaluation of the household's socioeconomic characteristics, captured by a uniform questionnaire, and the program is free of charge for the bottom two quintiles. Households in higher quintiles have to pay a yearly rate, increasing according to their socioeconomic status (for example, the current annual charge for a household in the fifth decile is around US$170). That said, according to the Department of Health, less than 1% of the affiliated households pay for coverage.

The insurance consist of four different modules of defined services that are administered free of charge. The first two, which have been in place since the start of the program, are the *Catálogo Universal de Servicios de Salud* (CAUSES), and the *Fondo de Protección contra Gastos Catastróficos* (FPGC). The former defines the principal set of covered interventions and medicines. At the program's start in 2004 it covered 91 interventions and 142 drugs (Ruvalcaba and Vargas 2010), by 2012 this had increased to 284 interventions and close to 400 drugs. The FPGC is a trust fund, jointly financed by the federal and state governments, that covers costly interventions which may result in catastrophic results (read, death or bankruptcy, or both) for the household affected. In 2012, the FPGC covered 58 interventions in areas such as cancer treatment, cardiovascular problems or organ transplants.

Of particular interest for the present analysis are the last two modules, implemented later. The first one is the *Seguro Médico para una Nueva Generación* (SMNG), which automatically covers every child born after December 1, 2006, and its immediate family with Seguro Popular. It also covers an additional 131(originally 110) interventions, directed at the principal causes of child mortality, during the child's first five years of life. Lastly, as a complement to the SMNG, in May 2008 the Department of Health implemented *Embarazo Saludable* (ES). It does not cover any additional interventions, as pre-natal care is already included in CAUSES, but automatically covers every pregnant woman and her immediate family with Seguro Popular.

In contrast to the public health care systems prevalent in the formal sector, the Seguro Popular was designed from its beginning primarily as an insurance, and not a service provider. As such, it buys services from a variety of health providers after prior certification through the state level bodies in charge of administering the program (called *Regímenes Estatales de Protección Social en Salud* (REPSS)). These service providers are in its majority health clinics run by the state level health departments, but in some cases, and especially for interventions through the FPGC, services can also be bought from other public or private providers (Ruvalcaba and Vargas 2010). At the same time, the states are expected, in coordination with the federal level, to direct health infrastructure investments to underserved areas, and 3% of the Seguro Popular budget is earmarked to finance such investments.

3 Data and Empirical Strategy

The data come from two principal sources: The micro sample of the 2010 Mexican Census, and monthly data on the number of individuals affiliated to the Seguro Popular in each municipality, provided by Mexican Health Ministry. The empirical analysis, therefore, consists of determining the effect of a treatment variable only observable at the municipal level on a rare outcome (infant mortality) at the individual level. The imprecise (yet unbiased) treatment variable, combined with the relatively low number of deaths observed in the data, make it difficult to find statistically significant relationships. Fortunately, the sheer size of the micro sample gives the analysis statistical power.

Mexico conducts a full population census every five years, called a *Censo* in years ending in zero, and a *Conteo* in years ending in five. The principal difference between the two exercises is that in years ending in zero a much larger number of characteristics is captured. Most importantly, the *Censo* administers an extended questionnaire, capturing a large number of individual and household characteristics, to roughly 10% of the Mexican population. The resulting micro sample contains more than 10 million individual observations.

Among other things, the extended questionnaire asks every woman 12 years of age or older for i) her total number of life births, ii) the number of children still alive, iii) the date of the last life birth and whether the child is still alive, and in case it has died iv) the age at death (in days, months and years). The time period under consideration is restricted to account for births starting with the program's roll-out and ending with births that occurred a year prior to being observed. Given that the data was collected during May and June 2010, the final sample consists of all life births observed between January 2004 and April 2009, yielding a total sample size of 836,809 observations.

Two different binary outcomes are of interest. I will first show results for infant mortality, i.e. deaths occurring within the first year of life. Subsequently, result will be presented when the outcome of interest is restricted to deaths occurring during the first month of life (neonatal mortality). Given the way the data are coded, I slightly diverge from the usual definition of neonatal mortality as death occurring during the first 28 days of life to a full calendar month. However, I do not believe this to be an important limitation of the interpretation of the results. For the infant mortality estimations, the sample needs to be restricted to births at least one year prior to data collection. For deaths occurring during the first month a larger sample could be used, however, for the sake of comparability of results it is preferable to work with a consistent sample across estimations.

3.1 Selection on outcome variables

The principal problem with the data is illustrated in table 1. As can easily been seen, the reported mortalities in the census micro data are consistently lower than the mortality estimates reported by UNICEF. Moreover, while the population mortality figures follow a downward trajectory, the mortality rates implied by the data suggest an increase. The root of the problem is that the census data contains a random sample of women, but not of births, as for each woman only the *last* live birth is observed.

(Table 1 about here)

This implies two different kinds of selection. First, the further back in time a women had a live birth, the higher is the likelihood that another birth occurred in the meantime, implying that the first birth is not observed. This likelihood can be expected to be a function of the time lapsed since the birth in question, the mother's age, and her number of surviving children at time of the birth. Since all these characteristics are easily observed in the data, this is a simple case of selection on observables and would not result in biased estimators. The second kind, however, is more of a problem. As most parents to some extent plan the ideal size of their families, a child that does not survive can be expected to be replaced by another birth. The probability of another birth is higher, the more time has passed since the death of the child in question. As a result, the probability of a non-surviving child to be in the sample is lower than that of a surviving one; and the gap between these probabilities will be larger for births further back in time. In addition, there is an issue of attrition due to incomplete data. A birth only enters the sample if year and month of occurrence are observed. If the child does not survive, year and month of death need also to be observed. It seems reasonable to assume that the respondent is more likely to have forgotten this information for a child who died than for one who survived. Again, this issue will be more severe the further back in time the birth occurred.

The patterns observed in table 1 clearly reflect these concerns, as true infant mortality figures are almost four times higher in 2004 than what is implied by the data. This gap is reduced to about twice the rate observed in the census data in 2009 (and would continue to narrow into 2010). For neonatal mortality as similar tendency, albeit with a somewhat narrower gap, holds. The upshot is a clear case of selection on the outcome variable, which, if unaddressed, would result in biased estimates.

3.2 Correction for selection

Feasible methods to correct for this kind of selection on outcomes can be found in the literature on choice-based sampling. Originally conceived as a method to address situations in which the sampling itself depends on the outcomes of interest[2], it can generally be applied in cases when the sampling probability is a function of the discrete outcomes of interest. In this paper, I will apply the weighted exogenous sampling maximum likelihood (WESML) estimator, proposed in the seminal paper by Manski and Lerman (1977). This estimator applies to cases in which the proportions of the different outcomes in the population are known, and has been shown by the authors to yield consistent estimates. That said, it is not the most efficient method available. For example, Amemiya and Vuong (1987) show that it is asymptotically less efficient than an alternative estimator proposed by Manski and McFadden (1981). Its most important virtue, however, is its easy implementation, given that all other available estimators (including Manski and McFadden (1981)) come at considerable computaional costs[3]. I believe this to be a trade-off worth taking, and that the WESML estimator provides the most useful method for policy analysis.

Following the original notation, let there be C different discrete outcomes of interest, and let i $\in$ C denote each of these outcomes. The WESML estimator consists of a weighted maximum likelihood estimation in which the

[2]The standard example used in most of the literature is the case of mode of transportation choice, where respondents are often sampled during their journeys at parking lots, train stations, bus stops etc..

[3]The interested reader is referred to Manski and McFadden (1981) and Cosslett (1981), who propose a number of different MLE estimators, many of which are computationally very intensive, under varying assumptions on the underlying data. Imbens (1992) proposes a method of moments estimator that is easier to implement than the ones proposed by Cosslett (1981), but is still computationally very costly. For a concise discussion on most of these estimators see Amemiya (1985), chapter 9.5.

weights applied to each observation n are defined as $w(i_n) = \frac{Q(i_n)}{H(i_n)}$; where $Q(i_n)$ denotes the population proportions of the outcome for observation n (which is assumed to be observable), and $H(i_n)$ denotes the same proportion in the sample. Given that in the present case these weights differ systematically over time, The weights will additionally be conditioned on the year of birth, t, as $w(i_n, t) = \frac{Q(i_n,t)}{H(i_n,t)}$[4]. More specifically, the UNICEF data presented in table 1 will be used as the observable population proportions in each year, $Q(i_n, t)$, whereas the mortality rates observed in the sample in each year will be $H(i_n, t)$.

3.3 Empirical specification and identification

The empirical strategy consists of a weighted probit estimation, with the outcome of interest being either death during the first year or first month of life, respectively. The principal interest is in the treatment variable capturing the extent of the program's roll-out in each municipality. In latent variable notation, the empirical model is:

[4]This modification is a straightforward extension of the original model. The proof for consistency can easily be modified to condition additionally on t. More specifically, expression (7) on page 1984 in Manski and Lerman (1977) would need to be changed to:

$$D_N(y,\theta) \overset{a.s.}{\to} D(\theta) \equiv \int_Z \sum_{i \in C} \sum_{t=1}^{T} \frac{P(i,t,z,\theta^*)H(i,t)}{\int_Z P(i,t,z,\theta^*)p(z)dz} w(i,t) logP(i,t,z,\theta)p(z)dz$$

$$\overset{a.s.}{=} \int_Z \sum_{i \in C} \sum_{t=1}^{T} P(i,t,z,\theta^*) logP(i,t,z,\theta)p(z)dz$$

$$
\begin{aligned}
y^*_{n,t,m} &= c + \alpha t_{m,t} + X_n\beta + \mu_m + \varepsilon_t + \epsilon_{n,t,m} \qquad (1)\\
y_{n,t,m} &= I[y^*_{n,t,m} > 0],
\end{aligned}
$$

where $y_{n,t,m}$ is the binary outcome for individual (i.e. birth) n, occurring in month t, and in municipality m. Furthermore, c denotes the constant term, $t_{m,t}$ is the treatment variable, and X_n represents a vector of time invariant, individual specific characteristics that will be included as additional control variables. The municipality and month (a binary variable for each month in the sample, not just calendar months) specific error terms μ_m and ε_t will be controlled for by fixed effects, and $\epsilon_{n,t,m}$ denotes the individual specific error term.

In order to arrive at unbiased estimates, in addition to the selection issues described above, the identifying assumption is that this last error term is uncorrelated with the treatment. Given that municipality and year fixed effects are controlled for, this boils down to the assumption that the program roll-out is not responsive to changes in municipality level characteristics that may also affect the outcomes. This difference-in-differences framework is standard practice in policy evaluation. Detailed municipality level characteristics are only available from Mexico's ten yearly census, making it infeasible to test whether or not program implementation was driven by changes in such observable characteristics. One can , however, test if implementation was driven by municipality characteristics in levels. If this is not the case, it seems fair to assume that changes did not have any effect either.

Table 2 shows results for a regression for the average monthly value of the treatment variable over the 2004-2009 period on a number of observable municipality level characteristics. Where possible, data from the year 2005 census (which captures fewer characteristics) was used. Other characteristics come from the 2010 census, but are unlikely to have changed significantly over the preceding five years or to have been influenced by the program. Lastly, each municipality's elevation in meters, its distance to the closest city with more than 100,000 inhabitants (in kilometers), and a measures of the mountainousness of its territory are included as truly time invariant characteristics. As explained above, Seguro Popular is a joint program between the federal and state governments. As a result, its implementation progressed at different speeds in different states. The second column in table 2, therefore, includes state-level fixed effects. As can be seen, once these are included almost all of the variables that are statistically significant in the first column become insignificant and have their point estimates reduced towards zero, implying that they mostly captured regional differences. The only characteristics that continue to be significant are, as one would expect, the proportion of the uninsured population and the total population in 2005.

(Table 2 about here)

The treatment variable is constructed based on data provided by Mexico's Department of Health, showing the number of individuals enrolled in the program on a monthly basis by municipality from its start in 2004 to late 2011. According to the Department of Health, universal coverage of the program was achieved in 2011. For that reason, I take September 2011, the last month in the data, to capture the total target population of the program, and calculate progress on a monthly basis as a fraction of that final number. It is important to note that Seguro Popular was not implemented at the municipality level, but rather locality by locality. Given that no data at the locality level is available (and even if it were, one is unable to observe the locality of residence for most households in the census micro sample), this is the best proxy for whether a household was exposed to the program one can come up with. Using the proportion of the population enrolled as a fraction of the target population, rather than the total population, avoids a measure that by construction is correlated with municipality level characteristics (as the target population is higher in more marginalized municipalities).

Two different treatment variables are constructed, one for roll-out in the month of birth and another one for the start of pregnancy, where the latter is simply the level of roll-out nine months prior (no information on premature births is observed). The objective is to test the importance of the program during pregnancy and after birth. The two measures will, of course, be highly correlated, making a clear-cut answer difficult, but some important information can be gained nonetheless. Given that the empirical specification controls for municipality level fixed effects, additional control variables are only introduced at the individual level (i.e. referring to the mother or the household). A wealth index is constructed by principal component analysis, using characteristics of the dwelling and information on a large number of household assets collected in the census. The other controls are the mother's age at birth (approximated by the mother's age in 2010, minus the years since the birth), its square, the number of previous live births, her completed years of schooling, whether or not she is indigenous , and whether she is currently working. All control variables were collected in 2010, while included births occurred from 2004 onwards. Therefore, some of them may not properly capture the characteristic in question at the time of birth. Some may even be affected by the outcome itself (e.g. years of schooling or whether a women is working). While the estimates on the control variables may have to be taken with a grain of salt, the aim of their inclusion is not to have a complete model of all factors affecting mortality risks, but rather to show that controlling for these factors do affect the estimates on the impact of the program itself.

Table 3 shows summary statistics for all relevant variables. The low mean values for infant and neonatal mortality reflect the numbers already shown in table 1. The average proportions of the target population covered by the program at the time of birth or pregnancy are 31% and 23%, respectively. Given some natural fluctuations in the size of the target (or total) populations and enrollment levels some of these figures are larger than one. A more detailed examination of the data shows that less than 5% of all municipalities have a value larger than one at some point in time, less than 2% have a value larger than 1.5 and only two municipalities have a value larger than 2. The last two entries in the table will be used in the robustness checks. They capture the per-capita number of health staff and health units in each municipality belonging to the Department of Health. These number are only available yearly, not monthly.

(Table 3 about here)

4 Results

The remaining tables show different sets of results. First, tables 4 and 5 show the primary results of interest, that is the program's estimated effect on infant and neonatal mortality rates,using the WESML estimator. After that, for the sake of comparison, table 6 shows the results that would have been obtained using a simple, unweighted probit model for the most important results. This will be followed by a series of robustness checks: Table 7 shows results for the inclusion of a number of additional control variables associated with the implementation of Seguro Popular. This is meant to shed some light on the question which components of the program had the strongest effects.

Lastly, table 8 addresses the possibility that the results in tables 4 and 5 may at least partly be driven by another source of selection not discussed so far: That the program may systematically affect the likelihood that a pregnancy results in a miscarriage or a stillbirth.

Given that all the tables show results for binary dependent variable models, the point estimates present expected marginal effects at the mean. Below, in parenthesis, p-values for statistical significance are shown. All tables also show the number of observations included, whether the sample is for the entire population or only the three lowest quintiles of the wealth index, and, where appropriate, whether the estimations are for neonatal or infant mortality. All results presented include municipality and monthly fixed effects.

4.1 Principal Results for Infant and Neonatal Mortality

Tables 4 and 5 show results for the entire sample in the first three columns, and those for the bottom three wealth quintiles in columns 4-6. For each sample, the first specification includes the two alternative treatment variables measuring program roll-out at the moment of estimated onset of pregnancy and at birth, respectively. This is compared in columns 2 and 5 to the effect if only the first treatment is included. Finally, in columns 3 and 6, results are shown for the inclusion of additional control variables.

Starting with columns 1 and 4, in both tables, and for both samples, the level of program roll-out at birth is statistically insignificant, and the point estimates are fairly small in magnitude. The effect at the onset of pregnancy, however, is consistently significant at the 5%-level (with the exception of column 1 in table 5, where it is slightly less) and much larger in magnitude. These results provide strong evidence that the program's main effect on survival take place during pregnancy. Moving one column to the right, the estimated effect at the start of pregnancy becomes statistically stronger and slightly larger in magnitude in all specifications. This is unsurprising, as the collinearity between the two treatments can be expected to increase standard errors if both are included. In columns 3 and 6, the inclusion of the additional control variables reduces the point estimates only very marginally, and does not affect the levels of statistical significance. This result provides additional evidence for the exogenous nature of the treatment variable and strengthens the validity of the results.

(Table 4 about here)

(Table 5 about here)

Focusing on this last case as the preferred specification, the estimates imply that a 100% roll-out of the program (as is claimed to have been achieved by late 2011) is expected to have reduced infant mortality rate by 4.94 deaths and neonatal mortality by 2.73 deaths per thousand live births on average over the time period under consideration. Comparing these numbers with total mortality rates presented in table 1, it may seem that almost the entire reduction in infant and neonatal mortality can be attributed to the program. It has to be kept in mind, however, that the last pregnancies in the sample are estimated to have started in July 2008. At that time, the average roll-out level per municipality stood at around 48%, implying that less than half the decrease is explained by the program. In addition, as shown in table 2, the roll-out was first targeted at municipalities with a high proportion of the population uncovered, mostly rural and poor places. These are also the areas where the program is likely to have had the largest effect on mortalities. The effect may therefore have been more muted as the program progressed. In that sense, the result presented here are unbiased estimates of the average effect of the program up to mid-2008.

Given that Seguro Popular is primarily aimed at households in the lower wealth quintiles, the primary interest should be in its effects on that segment of the population. As explained above, the program is free for households in the lowest two quintiles according to their socio-economic status. Here, results in the lowest three quintiles are shown for a variety of reasons. Firstly, given that all households with a child born after December 2006 can enroll in the program free of charge, the cut-off is insignificant for much of the sample. Next, the wealth index created is unlikely to be a perfect proxy for the much more detailed information used to determine actual status. Therefore, many households who are exempt from contributions may be above the second quintile in the data. Finally, by restricting the sample to households below a certain level of wealth, one loses additional observation from municipalities in which no deaths are observed in such households. Including the third quintile helps to minimize that loss. Moving to the results, as would be expected, point estimates increase in magnitude to an estimated reduction of 6.61 deaths for infant and 3.31 deaths out of a thousand live births for neonatal mortality.

While not of primary interest, it is worth pointing out that all the included control variables are highly statistically significant. As expected, higher wealth is associated with lower risk of mortality, even though this relationship is considerably weaker among only the bottom three wealth quintiles. There is also an inverted U-shaped relation between mother's age and the child's probability of survival. The risk of death is increasing in the number of a mother's previous birth, and decreasing in her level of schooling (this last result is potentially interesting given that household wealth is controlled for). Children born to indigenous mothers face a higher infant mortality rate of about 3 per one thousand live births (and about 2 for neonatal mortality). Lastly, there is a very strong positive correlation between a mother's labor force participation and her child's mortality risk, which, however, may partially be the result of reverse causation.

4.2 Comparison with Unweighted Probit Results

It is instructive to compare the results obtained with the WESML estimator, to those of the unweighted probit model. Table 6 does this for the preferred specifications with the start of pregnancy treatment variable and all the additional controls (corresponding to columns 3 and 6 in tables 4 and 5). Ex-ante, one would expect the unweighted model to underestimate the true effect, given that births of non-surviving children are systematically under-represented in the sample. A direct comparison with the results discussed above shows that, while similar in sign and significance, the unweighted probit model does indeed consistently underestimate the program's true effect by around 50-70%. Despite this huge difference in point estimates, their p-values are roughly similar. This is largely due to the loss of efficiency by using a weighted estimator. However, as argued above, the principal advantage of the WESML estimator is its easy implementation given that more efficient estimators that correct for the selection bias come at a very high computational cost.

(Table 6 about here)

4.3 Robustness checks

Tables 7 and 8 address two different concerns with the results presented thus far. The first one aims to shed light on the potential causal effects of different components of the program. As explained, program roll-out went hand in hand with an extension of health services provided by health departments. This raises the questions to what extent the observed results are driven by the simple presence of more health clinics and staff in the municipality, and not the free access to the services provided.

In order to tackle this questions, Table 7 shows results for the preferred specification (but the control variables are omitted in the table for ease of exposition) under the inclusion of additional control variables for investment in health services. Results are only shown for the sample consisting of mothers living in households in the lowest three wealth quintiles, as these are of primary concern, using the WESML estimator. The first four columns in table 7 show the effects if the per capita number of health clinics and health staff are included. Point estimates barely change with the inclusion of health units or health staff, nor do do levels of significance. The one exception to this is the effect of health staff per capita on the risk of neonatal mortality, which is significant at the 10% level. It also renders the direct effect of the insurance coverage insignificant at the 5% level. It is somewhat puzzling, however, that this effect disappears for the case of infant mortality. One possible explanation for this apparent contradiction is that the direct effect of health staff in this case is rather life prolonging than life saving.

(Table 7 about here)

Columns 5 and 6 in table 7 include dummy variables for the implementation of the program's extension to more pregnant mothers (*Embarazo Saludable*) in September 2009, and all newborn children (*Seguro Médico para una Nueva Generación*) in December 2006; plus their interaction term with the treatment variable of program roll-out. It has to be kept in mind that the dummy variables are simply equal to one after the date of implementation of the additional components. The variables of interest are, therefore, the interaction terms. No additional effect of these programs can be found, and their inclusion does not significantly alter the point estimates on the treatment variable (though it reduces the level of significance due to the induced additional collinearity).

Lastly, table 8 addresses the question whether the previous results may be biased by a selection effect due to the program's effect on pregnancy risks itself. This bias, at least in theory, may go either way. On the one hand, better pre-natal care may reduce the risk of a miscarriage or stillbirth, resulting in additional live births. However, if the surviving embryo or fetus faces a systematically higher mortality risk after birth, the results presented would actually underestimate the true effect of the program. On the other hand, better pre-natal care may also result in the early detection of medical problems in the embryo that may result in a low likelihood of survival. Since in such situations induced abortions are legal in Mexico, fewer such children may be born, resulting in an overestimation of the program's true effect.

(Table 8 about here)

The Mexican census does not collect data on pregnancies. For that reason this exercise will be done using a different dataset, the 2009 round *Encuesta Nacional de la Dinámica Demográfica* (ENADID). This survey contains a random sample of women aged 15-49 years of age, and collects detailed data on all pregnancies a woman ever had. While the data collected in this survey would in theory allow for the estimation of mortality risks without the selection concerns, its relatively small sample size makes this infeasible(only 502 death during the first year of life are observed of which 328 occurred during the first month). For the time period of interest, the survey has information on 37,230 pregnancies, of which 3706 (i.e. close to 10%) resulted in a miscarriage. Of these only 307 occurred after the third month of gestation (which can be used as a somewhat lose definition for a stillbirth). While it would be of interest to estimate the risk of a miscarriage at different months of gestation, the high concentration of miscarriages in the first three month renders such an exercise infeasible. Though would also be of interest to be able to distinguish between spontaneous abortions and abortions induced due to medical considerations. However, the data is not detailed enough to distinguish between the two. For that reason, the results in table 8 simply show the effect of the level of roll-out of Seguro Popular at the start of the pregnancy on the risk of a miscarriage.

Given that ENADID is also administered by INEGI, the questionnaire in some parts is very similar to to the one used in the Census and allows for the creation of comparable control variables. The only difference to the other results is that the wealth index is based on a smaller number of household assets, and that the number of previous pregnancies, rather than previous births is used. The first two columns in table 8 show results for the entire sample, and columns 3-4 only for the bottom three wealth quintiles. It can readily be confirmed that the program does not appear to have affected the risk of a miscarriage, as the sign of the estimate changes, between the two samples and in either case is highly insignificant. While the point estimates may appear comparable to those in the previous tables, it has to be kept in mind that the average risk of a pregnancy interruption is more than ten times higher than that of infant mortality.

One last potential source of selection bias that cannot be addressed comes from the unobserved maternal mortality during child births, as in both data sources the surveys sample women, not children. According to World Bank estimates this mortality amounts to 50 maternal deaths per 100,000 (or 0.5 per 1,000) live births in Mexico in 2010. Given this relatively low incidence compared to mortality risk, it seems fair to assume that its omission cannot have a large effect on mortality estimates.

5 Conclusions

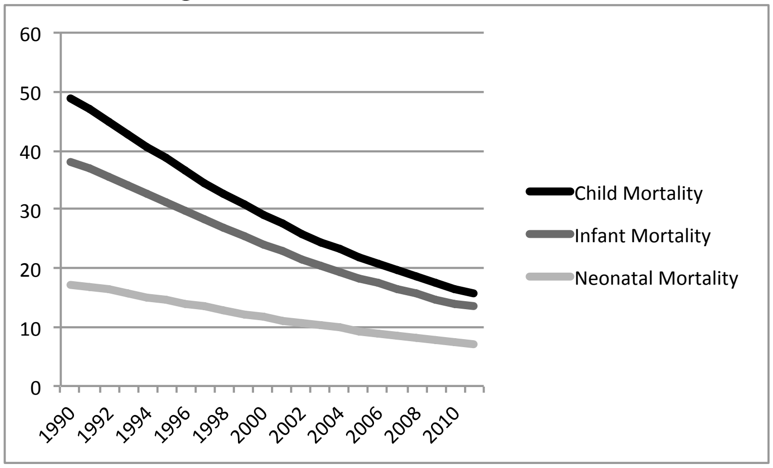
Evaluating the impact of access to medical services on health outcomes can be difficult in the short run, as, at least for adults, many such outcomes only change slowly over time. It is, therefore, a promising exercise to focus on impacts early in life. Not only would one expect a more direct effect, but outcomes such as child mortality are also among the most important public health targets. The main problem one encounters is finding sufficiently granular data that allow for the clear identification on a relatively rare outcome, all the more because detailed mortality statistics usually not available at the local level.

This paper assessed the effect of Mexico's non-contributory health insurance Seguro Popular on infant mortality over the first five years of program roll-out, using mothers' self-reported mortalities from the country's 2010 census micro data. This provided a big enough sample size, plus a number of household specific control variables. However, as is often the case with such data, it also provided data that is likely to be selected on the outcome variable.

Employing an estimation technique originally developed for the case of choice-based sampling, it was shown that Seguro Popular had a large and significantly negative effect on infant mortality. The risk of a child dying in the first month of life is reduced by close to 5 out of 1,000 (or 0.5%) for the population at large and by around 7 out of 1,000 (0.7%) for the program's target population. This appears mostly to be the result of the mother counting with health coverage during pregnancy, rather than medical attention during the year of life. Moreover, not taking into account the sample selection would have resulted in in large underestimation of the program's true effect, albeit with the same sign and levels of significance.

These results indicate that providing access to even fairly basic health care services can significantly improve health outcomes. It remains to be seen to what extent this is also true for outcomes that take longer to materialize and/or are affected by a less direct causal link, such as e.g. the incidence of cardiovascular problems. As time passes, future research should focus on the effects on such longer term outcomes.

Figure 1: Evolution of Mexico's under five mortalities.



Source: UNICEF

Table 1: Mortality rates in population and sample.

Year	**Infant Population**	**Infant Sample**	**Neonatal Population**	**Neonatal Sample**
2004	19.4	5.3	9.8	3.6
2005	18.4	6.1	9.4	3.9
2006	17.4	5.5	9.0	3.7
2007	16.5	5.7	8.5	3.7
2008	15.7	6.1	8.1	3.8
2009	14.8	7.2	7.8	4.8

Source: UNICEF

Table 2: Determinants of program roll-out

	(1)	(2)
Prop. Uninsured 2005	.116***	.247***
	(0.000)	(0.000)
Total Population	-6.63e-08***	-1.24e-07**
	(0.002)	(0.028)
Prop. Homes with Water 2005	.055**	-.014
	(0.013)	(0.637)
Prop. Homes with Sewer 2005	-.068	-.053
	(0.157)	(0.447)
Prop. Homes with Solid Floor 2005	.016	.009
	(0.459)	(0.794)
Prop. Homes with Electricity 2005	-.224***	-.003
	(0.000)	(0.967)
Prop. Illiterate 2010	-.331***	.069
	(0.000)	(0.178)
Prop. Economically Active 2010	.023	.021
	(0.719)	(0.813)
Prop. Secondary Education 2010	-.055	.155
	(0.296)	(0.135)
Prop. Indigenous 2010	-.038**	-.027
	(0.020)	(0.358)
Prop. Migrant Households 2010	-.0009	-.612
	(0.997)	(0.198)
Elevation	-.0000403***	-3.82e-06
	(0.000)	(0.722)
Distance Big City	-.0002**	-.0001
	(0.014)	(0.453)
Ruggedness	-.0002***	-.0000574
	(0.010)	(0.540)
Observations	2395	2395
State Fixed Effects	No	Yes

Notes: ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. P-values in parenthesis. Dependent variable is average level of program roll-out at municipality level over six year period. Robust standard errors.

Table 3: Summary Statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Infant Mortality	0.0059	0.0766	0	1	833886
Neonatal Mortality	0.0038	0.0616	0	1	833886
Insurance at Birth	0.31	0.27	0	3.79	833886
Insurance at Pregnancy	0.23	0.26	0	3.79	833886
Wealth	-0.22	2.59	-5.41	4.29	833351
Age	26.62	6.69	9	54	833886
Age Squared	753.21	380.49	81	2916	833886
Previous Births	1.78	1.96	0	23	833886
Schooling	7.68	4.15	0	24	828422
Indigenous	0.34	0.47	0	1	830722
Working	0.29	0.45	0	1	831491
Health Units at Pregnancy	0	0	0	0.01	801365
Health Staff at Pregnancy	0	0	0	0.02	773252

Table 4: Results for infant mortality

	(1)	(2)	(3)	(4)	(5)	(6)
Insurance at Pregnancy	-4.2**	-5.57***	-4.94***	-6.04**	-7.08***	-6.61***
	(0.048)	(0.002)	(0.005)	(0.015)	(0.001)	(0.001)
Insurance at Birth	-2.78			-2.13		
	(0.196)			(0.396)		
Wealth			-0.4***			-0.356*
			(0.007)			(0.070)
Age at Pregnancy			-1.31***			-1.2***
			(0.000)			(0.000)
Age Squared			0.03***			0.03***
			(0.000)			(0.000)
Previous Births			1.37***			1.43***
			(0.000)			(0.000)
Schooling			-0.62***			-0.58***
			(0.000)			(0.000)
Indigenous			2.81***			3.51***
			(0.000)			(0.000)
Working			3.81***			4.34***
			(0.000)			(0.000)
Observations	704458	704458	693500	538805	538805	531749
Wealth Quintiles	All	All	All	1-3	1-3	1-3

Notes: ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. Results for probit model with robust standard errors on binary variable indicating death in the first year of life. Point estimates show average marginal effect at the mean, p-values are in parenthesis. All specification include year-month and municipality level fixed effects.

Table 5: Results for neonatal mortality

	(1)	(2)	(3)	(4)	(5)	(6)
Insurance at Pregnancy	-2.71*	-3.23***	-2.73**	-3.44**	-3.77**	-3.31**
	(0.063)	(0.011)	(0.021)	(0.050)	(0.014)	(0.021)
Insurance at Birth	-1.04			-0.67		
	(0.479)			(0.707)		
Wealth			-0.26***			-.02
			(0.009)			(0.148)
Age at Pregnancy			-0.4***			-0.37**
			(0.007)			(0.040)
Age Squared			0.01***			0.02***
			(0.000)			(0.000)
Previous Births			0.66***			0.72***
			(0.000)			(0.000)
Schooling			-0.32***			-0.3***
			(0.000)			(0.000)
Indigenous			1.51***			2.21***
			(0.005)			(0.001)
Working			1.85***			2.1***
			(0.000)			(0.000)
Observations	628330	628330	618497	470607	470607	464494
Wealth Quintiles	All	All	All	1-3	1-3	1-3

Notes: ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. Results for probit model with robust standard errors on binary variable indicating death on the first day of life. Point estimates show average marginal effect at the mean, p-values are in parenthesis. All specification include year-month and municipality level fixed effects.

Table 6: Results for unweighted probit model

	(1)	(2)	(3)	(4)
Insurance at Pregnancy	-1.64***	-2.2***	-1.19**	-1.44**
	(0.007)	(0.002)	(0.021)	(0.022)
Wealth	-0.13***	-0.12*	-0.11**	-0.08
	(0.009)	(0.082)	(0.011)	(0.163)
Age at Pregnancy	-0.44***	-0.41***	-0.16**	-0.15*
	(0.000)	(0.000)	(0.012)	(0.051)
Age Squared	0.01***	0.01***	0.01***	0.01***
	(0.000)	(0.000)	(0.000)	(0.000)
Previous Births	0.46***	0.48***	0.28***	0.31***
	(0.000)	(0.000)	(0.000)	(0.000)
Schooling	-0.21***	-0.2***	-0.14***	-0.13***
	(0.000)	(0.000)	(0.000)	(0.000)
Indigenous	1.01***	1.23***	0.67***	0.96***
	(0.000)	(0.000)	(0.004)	(0.001)
Working	1.40***	1.60***	0.86***	0.98***
	(0.000)	(0.000)	(0.000)	(0.000)
Observations	693500	531749	618497	464494
Wealth Quintiles	All	1-3	All	1-3
Dependent Variable	Infant	Infant	Neonatal	Neonatal

Notes: ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. Results for probit model with robust standard errors on binary variable indicating death in the first month of life, conditional on having survived the first day. Point estimates show average marginal effect at the mean, p-values are in parenthesis. All specification include year-month and municipality level fixed effects.

Table 7: Results for robustness checks

	(1)	(2)	(3)	(4)	(5)	(6)
Insurance at Pregnancy	-6.81***	-6.43***	-3.17**	-2.73*	-3.05**	-5.33*
	(0.001)	(0.004)	(0.029)	(0.081)	(0.038)	(0.051)
Health Units at Pregnancy	-0.96		-0.92			
	(0.753)		(0.671)			
Health Staff at Pregnancy		0.23		-1.66*		
		(0.848)		(0.061)		
Embarazo Saludable					3.05*	
					(0.394)	
Embarazo*Insurance					-2.46	
					(0.421)	
Seguro Medico Nueva Generacion						3.08
						(0.505)
Seguro Medico*Insurance						0.09
						(0.975)
Observations	514595	488001	449989	424196	464949	354606
Wealth Quintiles	1-3	1-3	1-3	1-3	1-3	1-3
Dependent Variable	Infant	Infant	Neonatal	Neonatal	Neonatal	Infant

Notes: ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. Results for probit model with robust standard errors on binary variable indicating the respective mortality. Point estimates show average marginal effect at the mean, p-values are in parenthesis. All specification include year-month and municipality level fixed effects.

Table 8: Results for pregnancy interruption.

	(1)	(2)	(3)	(4)
Insurance at Pregnancy	5.88 (0.716)	2.1 (0.897)	-4.11 (0.818)	-7.22 (0.690)
Wealth		1.24 (0.219)		0.27 (0.815)
Age		-12.78*** (0.000)		-10.95*** (0.000)
Age Squared		0.32*** (0.000)		0.29*** (0.000)
Previous Pregnancies		-0.56 (0.688)		-2.52 (0.108)
Schooling		0.58 (0.269)		0.53 (0.419)
Indigenous		-15.57 (0.107)		-17.34* (0.068)
Working		22.56*** (0.000)		20.14*** (0.000)
Observations	34057	33195	23759	22937
Wealth Quintiles	All	All	1-3	1-3

Notes: ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. Results for probit model with robust standard errors on binary variable indicating whether a pregnancy did not result in a live birth. Point estimates show average marginal effect at the mean, p-values are in parenthesis. All specification include year-month and municipality level fixed effects.

References

AMEMIYA, T. (1985): *Advanced Econometrics*. Harvard University Press, Cambridge, MA.

AMEMIYA, T., AND Q. H. VUONG (1987): "A Comparison of Two Consistent Estimators in the Choice-Based Sampling Qualitative Response Model," *Econometrica*, 55(3), 699–702.

AZUARA, O., AND I. MARINESCU (2010): "Informality and the Expansion of Social Protection programs. Evidence from Mexico," unpublished working paper.

BARROS, R. (2008): "Wealthier But Not Much Healthier: Effects of a Health Insurance Program for the Poor in Mexico," unpublished working paper.

BOSCH, M., AND R. M. CAMPOS-VÁZQUEZ (2010): "The Trade-offs of Social Assistance Programs in the Labor Market: The Case of the "Seguro Popular" Program in Mexico," unpublished working paper.

CAMACHO, A., AND E. CONOVER (2013): "Effects of Subsidized Health Insurance on Newborn Health in a Developing Country," *Economic Development and Cultural Change*, 61(3), 633–658.

CAMACHO, A., E. CONOVER, AND A. HOYOS (2010): "Effects of Colombia's Social Protection System on Workers' Choice Between Formal and Informal Employment," unpublished working paper.

Cosslett, S. R. (1981): "Efficient Estimation of Discrete-Choice Models," in *Structural Analysis of Discrete Data with Econometric Applications*, ed. by C. F. Manski, and D. McFadden, pp. 51–111. Cambridge: MIT Press.

Duval-Hernández, R., and R. Smith-Ramírez (2011): "Informality and Seguro Popular under Segmented Labor Markets," unpublished working paper.

Gortmaker, S. L. (1979): "The effects of prenatal care upon the health of the newborn.," *American Journal of Public Health*, 69(7), 653–660.

Grogger, J., T. Arnold, A. S. León, A. Ome, and M. Triyana (2010): "Identificación y análisis de los efectos del Seguro Popular en el gasto en salud de los afiliados," Discussion paper, Centro de Investigación y Docencia Económicas, Evaluation commissioned by the Comisión Nacional de Protección Social en Salud (CNPPS).

Imbens, G. W. (1992): "An Efficient Method of Moments Estimator for Discrete Choice Models with Choice-based Sampling," *Econometrica*, 60(5), 1187–1214.

Levy, S. (2008): *Good Intentions, Bad Outcomes: Social Policy, Informality and Economic Growth in Mexico*. Brookings Institution Press, Washington, D.C.

Manski, C. F., and S. R. Lerman (1977): "The Estimation of Choice Probabilities from Choice Based Samples," *Econometrica*, 45(8), 1977–1988.

Manski, C. F., and D. McFadden (1981): "Alternative Estimators and Sample Designs for Discrete Choice Analysis," in *Structural Analysis of Discrete Data with Econometric Applications*, ed. by C. F. Manski, and D. McFadden, pp. 2–50. MIT Press Cambridge, MA.

Martínez, G., and N. Aguilera (2010): "The value of health and longevity in Mexico," unpublished working paper.

Muñoz Hernández, O., G. Cortes-Gallo, R. Pérez-Cuevas, Jasso-Gutiérrez, and J. Walsh (2010): "Evaluación de la morbilidad y mortalidad neonatales en recién nacidos afiliados al Sistema de Protección Social en Salud," Discussion paper, Centro de Estudios Econmicos y Sociales en Salud, Hospital Infantil de México Federico Gómez, Evaluation commissioned by the Comisión Nacional de Protección Social en Salud (CNPPS).

Ruvalcaba, L., and M. Vargas (2010): "Evolución del Sistema de Protección Social en Salud," Discussion paper, Centro de Investigación y Docencia Económicas, Evaluation commissioned by the Comisión Nacional de Protección Social en Salud (CNPPS).

Ruvalcaba, L. N., and S. Parker (2010): "Identificación y análisis de los efectos en las condiciones de salud de los afiliados al Seguro Popular," Discussion paper, Centro de Investigación y Docencia Económicas, Evaluation commissioned by the Comisión Nacional de Protección Social en Salud (CNPPS).

Scott, J., and N. Aguilera (2010): "Análisis de la factibilidad, alternativas de diseño y consecuencias distributivas de la construcción de un sistema integrado de salud en México," Discussion paper, Centro de Investigación y Docencia Económicas, Evaluation commissioned by the Comisión Nacional de Protección Social en Salud (CNPPS).

Shi, L., J. Macinko, B. Starfield, J. Xu, J. Regan, R. Politzer, and J. Wulu (2004): "Primary care, infant mortality, and low birth weight in the states of the USA," *Journal of Epidemiology and Community Health*, 58(5), 374–380.

Sosa-Rubi, S. G., O. Galarraga, and J. E. Harris (2009): "Heterogeneous impact of the Seguro Popular program on the utilization of obstetrical services in Mexico, 2001–2006: a multinomial probit model with a discrete endogenous variable," *Journal of Health Economics*, 28(1), 20–34.