The Effect of Basic Health Insurance for the Poor on Adult Mortality Risks: Mexico's Seguro Popular*

Tobias Pfutze
Florida International University

RESULTS STILL PRELIMINARY, DO NOT CITE!

Abstract

The single most important outcome when evaluating the performance of health care systems should be their effect on health status. This study assesses the effect of Mexico's public health insurance Seguro Popular on the mortality risks of adult beneficiaries. It finds a significant negative effect for males, but only small or no effects for females. Moreover, pre-existing health infrastructure is shown to be a strong mediating factor in the size if these impacts. The paper uses the universe of individual death registered in Mexico during the 2004-13 period, estimating discrete time survival models with two-sided truncation. Identification is achieved using the staggered roll-out of Seguro Popular at the locality level, controlling for locality, year, and age fixed effects.

JEL Classification: I13, I15, I18, O12
Keywords: Mortality, Health Insurance, Mexico, Right-Truncation

*Computing resources for this work were provided by Florida International University's High Performance Computing System.

1 Introduction

How best to provide universal access to health care services constitutes one of the most prominent public policy challenges not only in the developed world, but increasingly also in poorer countries. In particular, upper-middle income countries are in the process of extending these benefits to their uncovered population. One if the biggest such programs is Mexico's *Seguro Popular*, which today provides coverage to around half the country's population. While Mexico has thus archived close to universal health insurance, many open questions remain to be answered as to the quality and impact of the services provided. One question of particular interest is whether Seguro Popular had any significant impact on beneficiaries' health status, and, ultimately, their risk of mortality and their life expectancy. This is all the more important since international comparisons suggest that Mexico still has room for improvement on the last count. Even though life expectancy has increased greatly from less than 60 years in 1960 to 74.9 years (77.5 for women and 72.1 for men) in 2015, bringing it close to rich country standards, it is still the lowest value within the OECD. Moreover, there are considerable regional variations. In the rich northern state of Nuevo León, life expactancy is 76.4 years while in the poorest state, Chiapas, it is only 72.8 years.

This study shows that Seguro Popular had a significant negative effect on mortality risks for the adult beneficiary population. However, this effect is strongly concentrated on elderly males, while effects for females are much smaller. It is also found that these effects interact in an important manner with the pre-existing health care infrastructure. In urban localities (those with 15,000 inhabitants or more), program participation is estimated to reduce the mortality risk for a 70 year old man from 1.74% to 1.2% (a 31% risk reduction). For a similar woman, however, mortality risks decrease by only 3.6%. Moreover, certain pre-existing public health service provision is found to be a strong complement to program coverage in urban areas- for men and women alike. However, in semi-urban localities (those with 2,500-14,999 inhabitants), there is some weak evidence for a pre-existing health services to be a substitute.

The data are taken from publicly available mortality records over the 10-year period 2004-2013. Since individuals who survived this period are not observed, the data suffer, by construction, from right-truncation (in addition to the left truncation typical in survival data). This also means that survival models are the only feasible estimation method. The literature on estimating such models with right-(or two-sided) truncated data is surprisingly scarce. This issue is addressed with a discrete time version of the common Cox proportional hazard model that explicitly takes the truncation into account with a fairly straightforward method that was originally developed for the estimation of time to pregnancy. The treatment variable is constructed as the yearly level of coverage in each locality. The slow roll-out of the program, combined with locality, year and age specific fixed effects, allows for the identification of the treatment effect. The estimation strategy thus constitutes a difference in differences identification strategy applied to a survival model. This is to my knowledge the first study to estimate the effect of Seguro Popular on general mortality, and also the first one that applies this estimation method to mortality data.

There already exists quite an ample literature on Seguro Popular, and health care system in other middle income countries. These studies can be roughly subdivided into three branches. Following Santiago Levy's critique of non-contributory social protection programs (Levy 2008), the first one analyzes the effect of access to the program on beneficiaries' labor market choices. In particular, whether or nor it led to an increase in informality. Studies find either no effect or a relatively small shift towards informality ((Azuara & Marinescu 2010),(Barros 2008),(Bosch & Campos-Vázquez 2010),(Camacho, Conover & Hoyos 2010),(Martínez & Aguilera 2010)). Moreover, Conti, Ginja & Narita (2016) find not only a 2%-3% increase in informality, but also a reduction in wages in the informal sector. The second branch looks at how beneficiaries respond to the program in their usage of health services. The most important questions being whether they take advantage of the services offered, and if the program lowers the incidence of catastrophic health expenditures that could permanently increase poverty. On this count, Barros (2008) and Grogger, Arnold, León, Ome & Triyana (2010) find that Seguro Popular decreased out of pocket expenses. The latter also looks at catastrophic expenses, but finds an impact only in rural areas. On the other hand, Knox (2008) finds an impact on health care utilization, but not on spending. Gakidou, Lozano, Gonzalez-Pier, Abbott-Klafter, Barofsky, Bryson-Cahn, Feehan, Lee, Hernandez-Llamas & Murray (2006) and Knaul, Arreola-Ornelas, Méndez-Carniado, Bryson-Cahn, Barofsky, Maguire, Miranda & Sesma (2006) provide descriptive evidence that the program reduced catastrophic health expenditures and, in the former study, that it also led to an increased use of health services. Lastly, a systematic review of 49 quantitative studies by Giedion & Díaz (2010) finds evidence on both outcomes.

The third branch, into which the present study falls, looks at the effect on actual health outcomes. Only a few studies have found any important impacts. The just mentioned review of 49 studies concludes that there is no coherent evidence on that count. In particular, Barros (2008) does not find any impact of Seguro Popular on self-reported health status, nor hypertension. Similarly, Knox (2008), in a study covering the 2002-04 pilot phase, does not find any changes in self-reported health status, nor the ability to perform daily activities. Additionally, Duval-Hernández & Smith-Ramírez (2011) and the aformentioned work by King, Gakidou, Imai, Lakin, Moore, Nall, Ravishankar, Vargas, Tellez-Rojo, Hernandez-Avila, Hernandez-Avila & Hernandez-Llamas (2009) also fail to find any improvements in health outcomes. That said, several authors have pointed to various problems with such studies. Scott & Aguilera (2010) and Victoria & Peters (2009) point out the difficulty in trying to identify the effect of a recent policy change on slow moving targets such as most health outcomes.Giedion & Díaz (2010) and Giedion, Díaz, Alfonso & Savedoff (2009) attribute the absence of conclusive results on health outcomes to a lack of appropriate data. In addition, there are also a number of studies that do find health improvements. Ruvalcaba & Parker (2010) find a significant effect of Seguro Popular on the reduction in cholesterol, and, in some case, high blood pressure. However, no effect is found on the incidence of chronic diseases such as diabetes. The result on high blood pressure is confirmed by Bleich, Cutler, Adams, Lozano & Murray (2007) using propensity score matching on beneficiaries. In two related papers, Pfutze (2014) and Pfutze (2015) analyzes Seguro Popular's impact on infant mortality and the risk of miscarriage, respectively, using in both cases relative program roll-out between municipalities as the treatment variable. The first study shows a significant negative effect on the incidence of infant mortality, which can be attributed to coverage during pregnancy. The second paper finds similar results for the risk of miscarriage. Child mortality is also the subject of Conti & Ginja (2016) and Conti, Ginja & delValle (2017), who also find an important negative effect.

Universal health care provision policies have also been studied for a variety of other Latin American countries. In Costa Rica, emergency health services are provided free of charge to the uninsured population. Arguably for that reason, Cercone, Etoile, Pacheco-Jimenes & Briceno (2010) is not able to find any difference in utilization of health services or expenditures between the insured and uninsured populations. For the case of Peru, Bitrán, Muñoz & Prieto (2010) report that beneficiaries of the country's Integral Health Insurance (SIS) make more use of health services and have much lower levels of out-of-pocket expenditures compared to the uninsured. Giedion et al. (2009) find simiar results with respect to health service utilization for Colombia's subsidized regime. In Brazil in 2004, according to Barros, Santos & Bertoldi (2008), more than 80% of child births were paid for by the country's universal health service (SUS). Moving to health outcomes, Victoria, Aquino, doCarmo Leal, Monteiro, Barros & Szwarcwald (2011) find that a higher proportion of coverage of Brazil's Family Health Program (PSF)[1] is correlated with a smaller gap in infant mortality between the richest and poorest income quintiles. Also looking at newborns, Camacho & Conover (2013), using a regression discontinuity design on data from a single metropolitan area, find that the country's subsidized health insurance had a negative effect on low birth weight, and a positive one on the Apgar score. For Costa Rica, Cercone et al. (2010) find a better self-perceived health status among the insured compared to the uninsured.

The present study expands this literature by assessing the effects of public health insurance on mortality risks. It is striking that no other studies have taken a closer look at this outcome beyond child mortality. The likely reason for this gap can be found in the nature of the available data. While many household level surveys collect data on items such as health status, utilization of health services, health expenditures, or child mortality, they rarely capture information on deceased household members. The only reliable data on adult mortality, therefore, come from public records, and present a num-

[1] A program established in 1994 to bring health services to poor rural areas

ber of non-trivial econometric challenges. This study further contributes to the literature by showing how information from death certificates can be employed to estimate mortality risks. This should facilitate similar studies for other public health programs.

The next section describes Mexico's Seguro Popular health insurance program. This is followed by a detailed discussion of the methodology employed. After that, section four discusses the data used and presents descriptive statistics. Section five presents results, and finally, section six concludes and points to future research.

2 Description of the program

At the center of Mexico's public health system stands its payroll-financed health service provider for private sector workers, the Mexican Institute for Social Security, (*IMSS* by its Spanish acronym), founded in 1943. At the time, the underlying assumption was that economic growth would eventually bring the entire labor force into the formal sector of the economy, and hence into IMSS. Providing health coverage for the population not working under a formal employment contract was thus seen as a transitory problem, and not worth of much political attention. If at all, that population was served by a hotchpotch of federal and state level hospitals and clinics, mostly run directly by the corresponding health ministry. These facilities offered services of varying quality, only some of which were free of charge.

IMSS is not just a health insurance, but a health service that runs its own clinics and hospitals. While restricted to private sector employees, similar health systems exist for federal public employees (*ISSSTE*), for the public employees in the different states, the armed forces (*ISSFAM*), and the parastatal oil company Pemex. In addition, even though they are required to pay their payroll contributions to either IMSS or one of the other systems if they are employees, most members of the middle and upper-middle classes have a private health insurance (often provided by the employer as an additional benefit), or prefer paying out of pocket for private health services.

Official records of affiliation show that the expansion in the coverage of IMSS roughly corresponds to Mexico's period of high economic growth. That is, it stagnated starting in the 1970s. By the mid-2000s it covered only around one-third of the Mexican population, with a range of up to 50% of coverage in the richest states, and levels as low as 10% in the poorest (Scott & Aguilera 2010). Against this backdrop, the Mexican government decided in the early 2000s to start implementing a universal health insurance program with the aim to provide quality health care services to the hitherto uncovered population. The interested reader is referred to Frenk, Gonzalez-Pier, Gomez-Dantes, Lezana & Knaul (2006) for a detailed discussion of the motivations behind the program. Seguro Popular started its universal roll-out in 2004, following a pilot with randomized implementation at the locality level that had started in 2002. The program initially faced strong opposition not only from the established public health care providers (e.g. IMSS), but also from the Ministry of Finance given its projected large burden on the public finances (Lakin 2010). In order to ease this burden in the short run, and give the budget more time to absorb the additional expenditure, the 2004 law that officially created the program stipulated that only 14% of the eligible population could be enrolled in any given year, practically prescribing a seven-year implementation period. This staggered roll-out underpins this paper's empirical identification strategy.

Even though conceived as a partially contributory program, Seguro Popular is de-facto non-contributory. Eligibility to differentiated yearly fees is determined by a household's socioeconomic status, captured by a uniform questionnaire used to determine eligibility for all of Mexico's federal social programs[2]. The program is free for the bottom two quintiles, and for households above that threshold an increasing fee is charged, according the socioeconomic metric derived from the CUIS. In addition, starting on December 1, 2006, the program was extended by the *Seguro Médico para una Nueva Generación* (SMNG), later renamed *Seguro Médico Siglo XXI*, which provides free coverage to each child born after that date and its immediate family. This coverage was further extended in May 2008 by *Embarazo Saludable* (ES) to start with pregnancy. In the end, hardly anyone signed up for the program if coverage had to be paid for. According to the federal Health

[2]The Uniform Questionnaire on Socioeconomic Information (*CUIS*) is used for al social programs. However, unlike for example Colombia's Sisben, the information collected is not converted into a uniform index. Each program's administration decides how to use the information gathered independently.

Ministry, less than 1% of affiliated households are currently fee paying.

The program is co-financed and co-administered by the federal and the state governments, with the former paying the lion's share. For each affiliated individual a fixed yearly quota has to be paid into the insurance account, corresponding to roughly 12% of a yearly minimum wage. This is subdivided into $\frac{5}{6}$ paid by the federation and $\frac{1}{6}$ by the corresponding state. Unlike the payroll based public health services described above, Seguro Popular is a proper health *insurance*. That is, it pays the actual health service providers a pre-determined fee for each intervention, but does not provide any such services directly. This separation between payer and provider is meant to increase efficiency and reduce corruption (for example by overcharging patients or under-reporting provision) by providing an additional layer of checks and balances. The services covered, all of them free of charge to program beneficiaries, are determined by two modules: The *Catálogo Universal de Servicios de Salud* (CAUSES) and the *Fondo de Protección contra Gastos Catastróficos* (FPGC). The former stipulates the principal interventions and drugs covered. In 2004 it started with 91 interventions and 142 drugs. By 2016 this increased to 287 covered interventions and 647 drugs. The FPGC is a trust found, again jointly financed by federal and state governments, covering high cost interventions for medical emergencies. As of 2016, it covered 61 interventions. In addition, the aforementioned Seguro Médico Siglo XXI covers a further 149 interventions, related to child health risks, during the first five years of a child's life.

The providers health services can be bought from are certified at the state level by the *Régimenes Estatales de Protección Social en Salud*. Most of these providers are federal and state level clinics run by the corresponding Health Ministry. However, other public facilities (such as those affiliated with IMSS) and in some cases private providers (in particular for interventions associated with the FPGC) are also certified (Ruvalcaba & Vargas (2010), Lakin (2010)). Lastly, given that many areas are underserved by health facilities, 3% of Seguro Popular's budget is earmarked to finance investments in health service infrastructure. These investments are supposed to be carried out by the states in close coordination with the federal government.

3 Methodology

The principal data source, described in more detail below, are publicly available individual mortality records. These are then merged with locality level roll-out by year and other locality/municipality-level data. As mentioned, the principal econometric challenge is that by only observing the deceased, all observations are right-truncated (however, given that the event of interest is death due to natural causes, non-natural deaths will be treated as right-censored). While mortality records are a textbook example of right-truncated survival data, there is surprisingly little literature on how to deal with this issue, be it theoretical or applied. One solution, proposed by Gross & Catherine (1992) and which can be applied to discrete and continuous time models, consists of reversing time and thus converting the right-truncation into left-truncation. Yet, given that survival data often, and in particular the one used here, are also left-truncated, the method is of limited usefulness.

I follow the method proposed by Scheike & Keiding (2006) for discrete-time duration models with right-truncation. Left-truncation (i.e. delayed entry) can easily be accounted for with the standard method of using a memoryless distribution, in this case the standard log-log link function (which corresponds to the discrete time equivalent of the Cox proportional hazard model). In the discussion that follows, it is important to keep in mind that analysis time is an individual's age. To start, the probability without right-truncation that the duration (i.e. age of death), T is equal to t, for individual i is:

$$P(T_i = t) = \lambda_i(t) \prod_{j=E_i}^{t-1} (1 - \lambda_i(j)) = exp(-F_i(t-1)) - exp(-F_i(t)), \qquad (1)$$

where E_i denotes the time period of entry (i.e. age in 2004 if older than 20), and $\lambda_i(t) = 1 - exp((-exp(X_{it}\beta)))$ is the log-log hazard rate in analysis-time period t. Individuals that are younger than 20 in 2004 enter in the year of their 20th birthday. X denotes a vector of possibly time-dependent right-hand side variables. The expression $X_{it}\beta$ will be referred to as the linear form. From now on (and also following Scheike & Keiding (2006)) the term $F_i(t) = \sum_{j=E_i}^{t} exp(X_{ij}\beta)$ will be used, which as shown by the last term in equation 1 significantly simplifies the expression.

Introducing right truncation, the hazard function can now be written as:

$$\begin{aligned}
\lambda_i^T(t) = P(T_i = t | T_i \geq t, T_i \leq S_i) &= \frac{P(T_i = t)}{P(T_i \in [t, S_i)]} \\
&= \frac{exp(-F_i(t-1)) - exp(-F_i(t))}{exp(-F_i(t-1)) - exp(-F_i(S_i))} \\
&= 1 - \frac{exp(-F_i(t)) - exp(-F_i(S_i))}{exp(-F_i(t-1)) - exp(-F_i(S_i))},
\end{aligned} \tag{2}$$

where S_i denotes the time period truncation occurs. Note that in time period t=S the numerator of the last term in expression 2 is zero, so that the contribution to the likelihood function is one, and to the log-likelihood function zero. This means that deaths that occurred in 2013 do not contribute to the likelihood (do not offer any relevant information)) at their age of death, but only during the nine preceding years. This makes intuitive sense, since the probability of death in 2013, conditional on dying before 2014, is equal to one. Using the result in expression 2, the probability with right-truncation that the duration T is equal to t for individual i is:

$$
\begin{aligned}
&P(T_i = t | T_i \leq S_i) = \lambda_i^T(t) \prod_{j=E_i}^{t-1} (1 - \lambda_i^T(j)) \\
&= \left[1 - \frac{exp(-F_i(t)) - exp(-F_i(S_i))}{exp(-F_i(t-1)) - exp(-F_i(S_i))}\right] \prod_{j=E_i}^{t-1} \frac{exp(-F_i(j)) - exp(-F_i(S_i))}{exp(-F_i(j-1)) - exp(-F_i(S_i))} \\
&= \frac{[exp(-F_i(t-1)) - exp(-F_i(t))] \prod_{j=E_i}^{t-1} [exp(-F_i(j)) - exp(-F_i(S_i))]}{\prod_{j=E_i}^{t} [exp(-F_i(j-1)) - exp(-F_i(S_i))]}
\end{aligned}
\tag{3}
$$

taking logs:

$$
\begin{aligned}
log[P(T_i = t | T_i \leq S_i)] \; &= log[exp(-F_i(t-1)) - exp(-F_i(t))] \\
&+ \sum_{j=E_i}^{t-1} log[exp(-F_i(j)) - exp(-F_i(S_i))] \\
&- \sum_{j=E_i}^{t} log[exp(-F_i(j-1)) - exp(-F_i(S_i))]
\end{aligned}
$$

$$
= log[exp(-F_i(t-1)) - exp(-F_i(t))] - log[exp(-F_i(E_i-1)) - exp(-F_i(S_i))]
$$

$$
= log[exp(-F_i(t-1)) - exp(-F_i(t))] - log[1 - exp(-F_i(S_i))],
$$

where the last equation stems from the fact that $F_i(E_i - 1) = \sum_{j=E_i}^{E_i - 1} exp(X_{ij}\beta) = 0$.

The linear form includes year, age and locality-specific fixed effects. For the most parsimonious model it is:

$$X_{it}\beta = \beta_1 SP_{i,t} + \beta_2 Female_i + \theta_l + \lambda_t + \phi_y + \epsilon_{i,t}, \tag{4}$$

where i denotes each individual and t each analysis time period (i.e. age). $SP_{i,t}$ is the locality-level roll-out of Seguro Popular, and the most basic specification only controls for gender of individual i. θ_l are locality-specific, λ_t age-specific, and ϕ_y year-specific fixed effects. $\epsilon_{i,t}$ captures all unobservables. A separate fixed effect will be estimated for each included locality, rendering the inclusion of a general intercept pointless. In order to avoid collinearities, a baseline category is defined for the age and year fixed effects. For the former, this will be age 99, and for the latter the years 2012 and 2013. The need to use a two year baseline arises from the above discussed zero contribution to the log-likelihood from any observation that died in 2013. No separate fixed effects can be estimated for that year. However, a hazard rate corresponding to that year needs to be calculated for each observation in order to adjust for truncation. Assigning it the same value as the prior year is the best available option and only introduces a minimal inaccuracy in the results. Moreover, it should only slightly scale the estimated hazard rates, but have no impact on the estimated treatment effect.

Observations with a non-natural cause of death (homicides, suicides, and accidents), or who are 99 years of age, are treated as right-censored, but not right-truncated in the first case; and as both, right-censored and right-truncated, in the second. The former simply enter the likelihood every period with their normal log-log hazard functions $\lambda_i(t)$. For the latter, if their cause of death is natural, $\lambda_i^T(t)$ will be used (since they are also truncated), and the age specific fixed effect is assumed to being the same as for 99 year olds (mirroring the way that the fixed effects for 2012 and 2013 are treated as being the same). Their total contribution to the likelihood can then be expressed as: $P(T_i > O_i) = \prod_{j=E_i}^{O_i}(1-\lambda_i(j))$ (or $\lambda_i^T(j)$), respectively), where O_i is the age in the year before truncation occurs.

Following Scheike & Keiding (2006) and Fahrmeir & Tutz (2001), the estimation can be implemented using the Generalized Linear Model framework proposed by the latter, details will be discussed in the appendix. The likelihood function is maximized using using standard Newton-Raphson methods with a simple one one-sided line search. Starting with a step-size equal to one, it is cut in half at every iteration of the line search for as long as the value of the log-likelihood function continues to increase. The starting values for all parameters (including fixed effects) is set to zero. The convergence criterion is that between two iterations the proportional difference between any parameter value, fixed effect, and the log-likelihood is less than 0.0001 (or 0.01%).

4 Data

The principal data source employed in this study are the publicly accessible death certificates, made available online by Mexico's Ministry of Health.

Though anonymized, the data contain a rich set of information, such as the deceased's place of residence and, importantly, health insurance coverage. Mexico's national statistical institute (*INEGI* by its Spanish acronym) provides the same data in a more coherent, and preferable, format[3]. The data on program roll-out has been made available directly by the Seguro Popular administration (which forms part of the Ministry of Health). It consists of the number of beneficiary families in each locality at the end of each year. The size of each locality is taken from Mexico's 2005 census (called the *Conteo 2005*). The number of health staff at the municipal level, and the municipality population in 2000, are also available from INEGI.

The outcome of interest is the occurrence of a natural death. For that reason, observations whose cause of death is either accidental, homicide or suicide are treated as right-censored, entering the estimation up to the age in the year prior to their death. According to the Ministry of Health, complete program coverage was achieved in 2013. The treatment variable, measuring the progress of program roll-out in each year, is, therefore, constructed as the proportion of families enrolled in the program relative to the number in 2013. For cases in which a higher number was reported prior to that year, complete roll-out is assumed to have been achieved at that time, and the locality is assigned a value of one from that year onwards. Using the progress of roll-out relative to complete coverage, rather than coverage in

[3]The data provided by the Ministry of Health changes variable names and codification between different years. It also does not provide information at the locality level for all years

per-capita terms, avoids having to deal with a treatment variable that is partially a function of the initially uncovered population (which would raise additional concerns of endogeneity). The treatment variable must thus be interpreted as the probability that a person in the target population is a Seguro Popular beneficiary.

It is important to note that individuals are only observed at the time of their death. The data are then put into a dynamic format under the assumption that all the observed characteristics, in particular locality of residence, did not change since 2004. The target population consists of individuals who at the time of their death did either not have any health insurance, or were Seguro Popular beneficiaries. Even though program participation is observed in the year of death, it is not used as a treatment since it is unknown in which year the individual entered the program. Furthermore, using actual beneficiary status would likely raise endogeneity issues due to selection into the program, which at the locality level are taken care of by the fixed effects[4].

The final dataset consists of registered deaths between 2004 and 2013 of persons aged between 20 and 99, who either had no insurance or were Seguro Popular beneficiaries at the time of their passing, and who resided in localities larger than 2,500 inhabitants. The dataset is restricted to observations whose death was registered in the year of occurrence. This eliminates only a

[4]According to the data, 55.92% of the deceased in the target population were covered by Seguro Popular in 2013, the year that complete coverage was supposedly achieved. This percentage does not differ much by age group. The most likely implication is that in many cases beneficiary status was not properly reported.

very small number of observations (less than 1%). However, keeping deaths registered in later years would result in an unbalanced sample, as it would artificially increase the number of observations who passed in earlier years. Additionally, some observations are lost due to incomplete information in any of the other variables of interest, eliminating another 2%. The biggest attrition comes from the inclusion of pre-program health staff at the municipal level, which has missing values for several municipalities. As can be seen from table 3, this would reduce the number of spells by more than 200,000 or 15%. In order to avoid this large attrition, only the estimations that include the interaction with pre-existing health facilities will be run on the smaller sample. For the sake of comparison the corresponding baseline results for the smaller sample will also be presented.

The year 2004 is chosen as the starting date in order to exclude the pilot period in 2002/03. While the assignment into treatment was random, the selection into the pilot was not (though the localities selected supposedly constituted a representative sample at the national level). Moreover, while the number of participating families is observed at the end of 2003, and it is known which states entered in each of the two years, there is no data on the actual number of beneficiaries at the end of 2002. The restrictions to localities of 2,500 inhabitants or more, as well as the further division at 15,000, follows the official classification used by INEGI into rural (less than 2,500), semi-urban (2,500 to 14,999), and urban (15,000 or more) localities. Rural localities, some of which can be as small as a couple of households, are excluded because of the impossibility to separately estimate all their fixed effects, given the small number of observed deaths in most of them. For the same reason, I additionally imposed the restriction that a locality must have more than five observations and that at least one of them must be non-censored (i.e. be either a non-natural death, die at age 100 or over, or die in 2013) and that not all of them die upon entering the sample (either in 2004 or at age 20), so that at least one zero outcome is observed. This restriction only removes a total of 87 spells from the semi-urban localities (and none in urban ones), which corresponds to 0.008% of all spells and 0.026% of those in semi-urban localities.

Furthermore, there is a well-known, yet under-researched, problem that not all deaths occurring in Mexico are properly registered[5]. Though there is no hard data, this problem is very likely concentrated in small, rural communities due to the simple fact that in more urban setting one cannot bury or cremate a body without the need to produce some official documents. The upshot is twofold: Firstly, there would be even fewer observations in many small rural localities. Secondly, if Seguro Popular improves registration (which is likely), the results will be positively biased. To the extent that sub-registration of deaths is still an issue in semi-urban and urban localities, the latter point will still be a caveat to keep in mind when interpreting the results presented below. However, it has to be stressed that it would induce a positive bias, and thus work against the results found (which would then

[5]for a brief discussion see Gómez-Morín-Escalante (2015)

be under-estimates of the true effect).

The restrictions on age are imposed for two reasons: At the lower end to exclude minors, and at the upper end to exclude the few cases of extreme longevity. The principal aim of this study is to assess the impact of Seguro Popular on adult mortality. As discussed above, other studies have already established its effect on child mortality. Also, since child mortality has different causes from the adult sort, its inclusion would complicate the interpretation of the results. This last point is also true for children older than five, for who, additionally, mortality is very low. The restriction is implemented by assuming that an observation becomes of risk at age 20. For example, a person who died in 2010 at age 23 would only enter the estimation in 2007. This restriction may pose a slight risk of selection at the lower end, since Seguro Popular may affect the probability of surviving to age 20. However, the low risk of mortality at that age makes this concern negligible[6]. At the upper end, observations are censored at age 99 (e.g. someone who died in 2010 at age 102 would be observed from 2004 to 2007). Since age is also controlled for by fixed effects, including centenarians would at some point pose estimation problems due to the small number of observations at such an advanced age. Also, as discussed above, observations who die in 2013 do not contribute any information to the maximum likelihood function in that year (since the probability of death, conditional on being observed, is equal

[6]There is one more reason for entry at age 20. The youngest observations included were eleven years old in 2004. If the age of entry was lowered to, for example, ten years of age, the presence of child mortality would make selection more of a concern.

to one), but provide information up to 2012.

Figure 1 shows the evolution over the ten years under observation of the total number of deaths used in the analysis. It does so for three different groups: i) The population of interest, ii) the population that has some other form of insurance (either in one of the formal sector payroll-financed system or a private insurance), and iii) the population for who insurance status in unknown. It can be observed that the population of interest follows relatively closely the trend of the population with some other form of insurance, while the population with unknown insurance type stays roughly constant over time. The concern here is that if if Seguro Popular increases the number of deaths registered in the target population, the resulting estimates would be upward biased. Yet, if it resulted in beneficiaries being wrongful registered as IMSS members (or of some other insurance), the results would be downward biased (i.e. over-estimating any negative effect on mortality risk). In order to shed more light on this selection concern, table 1 runs a simple fixed effects regression at the locality level. The outcomes are the number of reported deaths in the target population in each locality (with more than 2,500 inhabitants) per year. The regression hence controls for locality and year fixed effects. The treatment variable is the same as used throughout. There is no significant effect on the average number of deaths reported in the target population, except for a small decrease in semi-urban localities. The latter may result in a small negative bias in the estimates shown below. However, the effect is not only barely significant, but also small in magnitude.

No such concern exists for urban localities.

The next two tables provide summary statistics on the different populations of interest (the full sample, and the division into semi-urban and urban localities). Table 2 shows the level of roll-out and the number of observed natural deaths over the 2004-2013 time period. Regarding roll-out, there is no big difference between the semi-urban and urban localities, though in the former the process was slightly faster. Given the estimation strategy, one has to exclude individuals that were either younger than 20 years (or 20 years old and died in 2013), or 100 years or older in 2004, or died of non-natural causes at age 20 or in 2004. Overall, there is a very similar increase over time with some drops in 2006 and 2011 which are likely due to year-specific shocks in data collection that should be taken care of by the year fixed effects and would not be expected to be related to program roll-out. The increase in deaths over time is only to be expected. For one, registration may have improved over these ten years. But arguably more important, almost all of the observed deaths occurred in cohorts that were born in years of very high population growth. As these cohorts enter ages of higher mortality risk, the number of deaths naturally increases.

Table 3 shows summary statistics for all the variables of importance in the estimations. Mostly self-explanatory, some points need further discussion. Firstly, the higher average age at death in semi-urban localities must not be interpreted as a higher life expectancy, but rather reflects different demographic characteristics. Smaller localities tend to have an older population as the younger migrate out. This is also reflected in the higher proportion of included spells that were younger than 20 years in 2004, and hence enter the estimation at a later date (again, it is assumed that an individual becomes of risk at age 20). The same can be said of observations that die at age 100 or older, of which there are slightly more in semi-urban localities. The (right-censored) non-natural deaths (homicides, suicides and accidents) make up less than 9% in semi-urban localities and more than 11% in urban ones, with an overall average of about 11% . The higher life expectancy of females is represented by their lower share in the data (less than 50%), but is driven by their lower incidence of non-natural deaths. Though omitted from the table, about 20% of male deaths are due to non-natural causes, while this is only the case for about 5% of female deaths. The average observation lives in a fairly big city with more than 300,000 inhabitants. Around 60% of observed deaths occur at age 60 or older. This may seem low, but one needs to take into account the much larger population of under-60 year olds in Mexico. Again, the difference between the two types of localities is explained by different demographics. Lastly, the two variables for health staff at the municipal level have been demeaned in order to provide a clearer interpretation of the parameters on interaction terms presented below.

5 Results

Results are presented in tables 4-9. It is important to note that the tables with estimation results present parameter estimates that cannot be directly translated into marginal effects. However, they allow for the assessment of statistical significance and relative magnitude of the effects. For ease of interpretation, only estimates on the variables of interest (those that assess the effect of Seguro Popular) are included, but all estimations include a full set of fixed effects (locality, age, and year), a control for gender, and all the necessary interaction terms. There are a number of statistics at the bottom of each table, showing the number of spells, the number of localities included, The log likelihood after convergence, and the number of iterations needed for the Newton-Raphson algorithm to converge.

5.1 Principal results

Table 4 shows result for the most basic specification, and compares them to the population covered by other, either public payroll-based or private, insurance schemes. The effect of Seguro Popular on mortality risks is statistically highly significant in the target population (the first three columns), with a somewhat lower effect in the semi-urban localities compared to the urban ones, where the parameter value more than doubles.

One direct way to assess the validity of the estimation approach taken here is to contrast it with the estimated effect on the population already covered by some other form of health insurance. This is done in columns 4 to 6 of the same table. A negative, yet much smaller, effect is found in the full sample. This effect is driven by a statistically significant negative effect in urban localities, with no effect in semi-urban ones. The effect on the former ones is, however much smaller (around one-third of the size) than the one found for the target population. These results largely support the approach taken in this paper, but also raise some interesting questions. The negative effect in urban localities could simply be due to relatives of a deceased person misreporting Seguro Popular as some other health insurance. Another possible explanation is that some people receive Seguro Popular even though they are already covered by some other health insurance. A third possibility is that higher coverage of Seguro Popular was accompanied with an improvement in publicly provided health services, thus acting like a positive externality.

Results become more heterogeneous when the program's effect is conditioned on age and gender. This is done in table 5 by interacting the treatment with a binary variable for each gender and with age, respectively. Table 6 then combines the two. Since age enters the estimation as a battery of fixed effects, the interaction terms employed here only capture the difference in the effect of Seguro Popular when the sample is divided into person younger than 60 years of age, and 60 or older. This cutoff roughly corresponds to dividing the number of deaths observed in equal parts (keeping in mind that the younger cohorts are much larger). The first three columns show that the effect is consistently much larger for males than for females, roughly by a factor of four to five. However it is still statistically significant for females.

When interacted with age, the program's effect is concentrated on the older population. Only in urban localities is it possible to identify an effect on the younger cohorts. Splitting this interaction further, by allowing for a separate interaction term with age for males and females, table 6 further strengthens the results on males. The effect is found to be statistically significant for older, as well as, younger males, with the effect on the former being larger by a factor between two and three. The effect on females loses significance in semi-urban localities, but is still statistically significant, and of similar magnitude, in urban areas (albeit only at the 10% level for the younger cohorts). The principal take-away from these estimations is that the program principally benefitted males, in particular older ones, while it had a much smaller effect on females.

5.2 Results in terms of mortality risks

The tables just discussed present results in terms of the parameters on the regressors in the linear form of the log-log link function. It is of inherent interest to translate these into estimates of actual mortality risks. In discussing these hazards, it is important to keep in mind that they refer to the age-specific (yearly) risk of a natural death, and not of global mortality. As a first step, figure 2 plots the estimated baseline hazards for the target population, based on the model in column 1 of table 4. The age specific hazards are computed as the average predicted risk in the average locality in 2012. The risk of a natural death at age 20 is estimated as 0.52%. It then increases to reach 1% at age 50, and 2% at age 77, and moves close to 20% when approaching age 100. It can be seen that the risk increases linearly until around age 70, and then starts to increase in a more exponential fashion. The increase in the estimated hazards is almost continuous. The exceptions are the higher reporting of round ages (in particular 30, 50, 60, and 70), which result in small spikes, and a spike at age 98. This last spike may be the result of only a small number of observations reaching that age, or many deaths occurring at age 99 being recorded as age 100.

Table 7 translates the results on the parameter estimates from table 6 into differences in risks of natural death. Hazard rates are reported for 40 and 70 year old males and females, for each of the three groups of localities. One result that immediately stands out is that females face higher risks than males even before the implementation of Seguro Popular, this is particularly pronounced for relatively younger women (40 years of age) than older ones (70 years). However, it has to be stressed that this only reflects risk of death due to natural reasons (it would be lower for women if homicides, suicides, and accidents were to be taken into account), and for the population that has no other health insurance.While at first surprising, this result may, together with the finding that Seguro Popular mostly benefitted males, point to a broader problem with Mexico's pre-existing health care system. It is beyond the scope of this paper to analyze whether the problem resides in the actual provision of health care services, or deeper social and/or cultural factors that leave women's' health needs unattended. The estimated marginal effects of the program for males are fairly high. The reduction of the mortality risk by 0.44 percentage points for all males ahed 70, corresponds to a 27% reduction in the risk (20% for semi-urban localities, and 31% for urban ones). For 40 year old males, the risk is still reduced by 14% overall.

5.3 Interaction with pre-existing health facilities

One important, and under researched, question is the interaction of health insurance and health services. As discussed, the program roll-out was accompanied by investments in health facilities in under-served areas. Unfortunately, the federal health ministry is not involved in the certification of providers, and does not keep any centralized data base on them. Moreover, a detailed analysis of their selection and impact would be beyond the scope of this paper. However, it is important that most of these facilities are run by federal or state health ministries; and that, where present, pre-existing public health facilities were certified. Some data on public health facilities, number of medical staff and number of units, is readily available at the municipal level from INEGI. Since the number of units can cover anything from a small clinic to a large hospital, the focus here will be on staffing levels. In the following tables, results without the interaction term for the corresponding smaller sample have been included for ease of comparison.

The numbers include doctors and nurses, and are not available at the locality level. Though far from perfect, they should provide a reasonable proxy for local coverage. I distinguish between staff in facilities belonging to the health ministries (either federal or state) and other public facilities (mostly IMSS). The data are available on a yearly basis, yet there is little cross-sectional variance in changes over time. For this reason, only the staffing levels of 2002, right before the start of the pilot phase, are used. The results have thus to be interpreted as the interaction of insurance coverage with pre-existing health services. This should be of interest in and of itself, and not just be seen as a proxy for the effects of health service expansion. As explained in the summary statistics in table 3, the number of staff is put into per-capita terms and, for ease of interpretation, demeaned. Given that most municipalities only have one semi-urban or urban locality, the per-capita medical staff levels in 2002 will be captured by the locality fixed effects, so that only the interaction term with Seguro Popular coverage is added to the model.

Table 8 shows the results from this exercise. The first three columns show results for the interaction with per-capita staff at facilities run by the corresponding health ministries. The other three the interaction with staffing levels at other public health facilities. The results are twofold: Firstly, pre-existing staff in Health Ministry facilities does not seem to matter at all for the effect of the subsequent program implementation. Yet, the interaction with pre-existing health staff at other public facilities is statistically significant. Secondly, and as an extension to the last point, the sign of this latter effect differs in semi-urban and urban localities. In the first, the effect is smaller in its absolute magnitude, yet statistical significance, and positive. This implies that Seguro Popular and health facilities act as substitutes. In urban localities, however, the parameter on the interaction term is negative (and larger in magnitude and significance), implying that the two are complements. The magnitude implies that an increase of two standard deviations above the mean in medical staff per capita would double the effect of Seguro Popular coverage. Table 9 further divides the analysis into the effects on men and women, corresponding to the models in the first three columns of table 5. In line with the previous results, staff at health ministry run facilities have no effect. The effect of non-health ministry staff is also similar, with only a few marginal differences. The interaction term in semi-urban localities becomes mostly statistically insignificant, though does not change much in magnitude. More interestingly, even though the effect of Seguro Popular at the mean is not statistically significant for females in urban areas, the interaction term with staff at other health facilities is. Though smaller in magnitude than the one for men, it implies that any effect of Seguro Popular on women in urban localities is mediated by the these pre-existing facilities.

These results have several implications. First and foremost, they do not bode well for the quality of care provided in Health Ministry run hospitals and clinics. This should be taken into account when directing investments to such facilities. A possible explanation for this result are differences in the quality of the services provided. More intriguing is the difference in the interaction term between coverage and other pre-existing public facilities. For example, one may assumes that those in larger localities provide higher quality services or more specialized and/or advanced treatments. Given that some people who are enrolled in Seguro Popular already have coverage from some other public insurance, the positive interaction term may simply indicate a duplication of services. In urban localities, however, the effect of Seguro Popular could be enhanced by providing access to higher quality care.

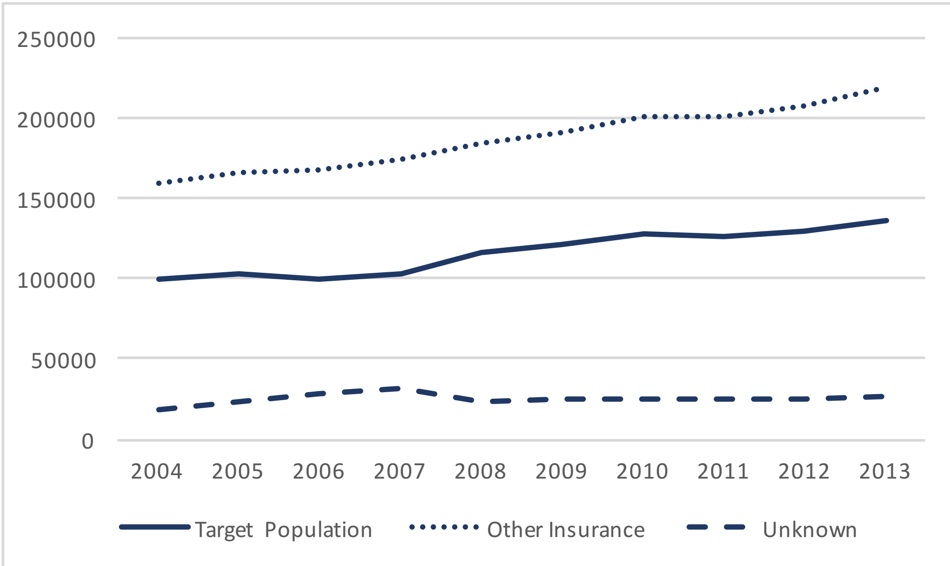
6 Conclusions

The effect of publicly provided health insurance on health outcomes is a research area of obvious importance. Mexico's ambitious Seguro Popular, which provides basic health coverage to around half the country's population, is of particular interest given its potential to provide a blueprint for other middle-income countries. One problem in reliably assessing its effects lies in obtaining systematic health data other than self-reported outcomes. This paper addresses the question by estimating the Seguro Popular's effect on mortality risks among the adult population, using information on all registered deaths in Mexico during 2004-13. The treatment variable is roll-out at the locality level (for localities with at least 2,500 inhabitants) in a difference in differences framework that accounts for locality level fixed effects. Mortality has the advantage that, beyond its obvious intrinsic interest, it is easily observable and captured by administrative data sources. The only practical problem arises because, in the absence of survey data that captures detailed mortality statistics, at each point in time the surviving population is not observed, resulting in right-truncated data. This problem is addressed by directly accounting for the truncation in a discrete time hazard model.

The results presented in this study allow for three clear-cut conclusions: i) Seguro Popular significantly reduced health risk for the adult male population, but ii) it had a much smaller effect for women. Moreover, iii) the effect on men is larger for the older cohorts. Additionally, the effect seems to be larger in urban than in semi-urban localities. This last result needs to be taken with a number of important caveats. For one, the program seems to also have had a small, yet significant, negative effect on individuals covered by some other insurance. Neither can it be ruled out that the rate of non-registration of deaths is larger in semi-urban than in urban localities. If Seguro Popular improved registration, this would put a positive bias on the results in the former. That said, the study also shows evidence that certain pre-existing health services, measured by per-capita staffing levels, are an important complement to insurance cover for beneficiaries in urban localities. Since it can be expected that larger localities have higher quality and more specialized health services, it seems reasonable that the larger effect of Seguro Popular in such places is driven by pre-existing facilities.

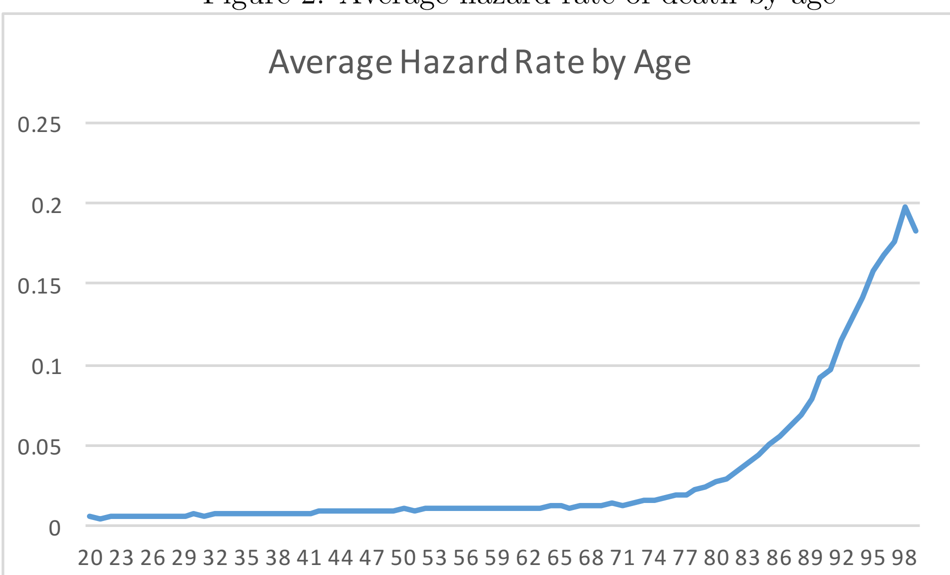
The most troubling result from a policy perspective is the much smaller effect for female beneficiaries. Moreover, the baseline mortality risks, for a natural death, estimated in this study are also higher for females than for males. Taken together, these point to a strong gender bias in Mexico's health care system for the poorer population. The precise reasons, which may lie in the system itself or merely reflect other social or cultural factors, are beyond the scope of this study, but definitely merit closer examination. Somewhat related, another area that should be the focus of future research are the program's effects on particular health risks and causes of mortality. Lastly, the interaction of insurance and health service quality constitutes another wide open research question. The results presented here point to a strongly complimentary effect, and also imply that the quality of health services accessible through the program should get a closer look.

Figure 1: Number of included deaths by year and insurance status.



Note: Total number of registered deaths in each year. Excluded are deaths of non-citizens, that did not occur in the year of registration, and for which no age is observed.

Figure 2: Average hazard rate of death by age



Note: Hazard rates estimated based baseline assumptions (year 2012 and no Seguro POpular benefit) on all included localities and assigning their average fixed effect value.

Table 1: Effect of Seguro Popular roll-out on number of registered deaths in the target population.

	>2.5k	Semi-Urban	>15k
Seguro Popular	1.174	-.596*	6.745
	(1.015)	(.351)	(7.094)
Num. Obs.	30603	25215	5388
F statistic	30.147	110.487	17.237
Mean of Outcome	45.69	15.26	188.11

Note: Dependent variable is the number of registered deaths in the target population in each locality/year. All estimations include locality and year fixed effects.

Table 2: Program roll-out at locality level

	Localities >2,500		Semi-Urban		Localities >15,000	
year	Roll-Out	Observed Deaths	Roll-Out	Observed Deaths	Roll-Out	Observed Deaths
2004	0.076471	99,897	0.0773201	28,106	0.0723487	71,791
2005	0.1415841	103,500	0.1438053	29,012	0.1308005	74,488
2006	0.2514989	99,729	0.2556072	26,498	0.2315543	73,231
2007	0.3529394	102,347	0.3570563	26,967	0.3329525	75,380
2008	0.4319053	116,229	0.4376884	33,190	0.4038298	83,039
2009	0.4856132	121,474	0.4921562	34,578	0.4538481	86,896
2010	0.6969964	128,132	0.7035037	37,061	0.665405	91,071
2011	0.8570722	126,048	0.8579978	36,418	0.8525783	89,630
2012	0.9161596	129,014	0.9150792	36,422	0.9214043	92,592
2013	1	135,565	1	38,246	1	97,319

Table 3: Summary Statistics

	Localities >2,500			Semi-Urban			Localities >15,000		
	Obs.	Mean	Std.Dev.	Obs.	Mean	Std.Dev.	Obs.	Mean	Std.Dev.
Seguro Popular	1,302,512	0.2668	0.4423	357,772	0.3070	0.4612	944,740	0.2516	0.4339
Age at Death	1,302,512	64.44	20.21	357,772	67.33	19.39	944,740	63.34	20.40
Younger than 20 in 2004	1,302,512	0.0282	0.1656	357,772	0.0213	0.1444	944,740	0.0308	0.1729
Death at 100 or Older	1,302,512	0.0069	0.0826	357,772	0.0077	0.0875	944,740	0.0065	0.0806
Violent Death	1,302,512	0.1079	0.3103	357,772	0.0874	0.2824	944,740	0.1157	0.3199
Female	1,302,512	0.4408	0.4965	357,772	0.4491	0.4974	944,740	0.4376	0.4961
Population	1,302,512	383,952	504,503	357,772	6,940	3,631	944,740	526726	526017
Older 60 at Death	1,302,512	0.6091	0.4879	357,772	0.6778	0.4673	944,740	0.5832	0.4930
Ministry Health Staff 2001	1,082,180	0.0000	0.0004	322,111	0.0000	0.0005	760,069	0.0000	0.0004
Other Public Health Staff 2001	1,082,180	0.0004	0.0008	322,111	0.0000	0.0005	760,069	0.0003	0.0008

Table 4: Main results

	Target Population			Other Insurance		
	>2.5k	Semi-Urban	>15k	>2.5k	Semi-Urban	>15k
Seguro Popular	-0.1432***	-0.0791***	-0.172***	-0.0503***	-0.0037	-0.0617***
	(-14.67)	(-5.56)	(-11.97)	(-5.44)	(-0.20)	(-5.56)
Number Spells	1,302,512	357,772	944,740	1,975,079	227,907	1,747,172
Number Localities	3,090	2,547	543	2,972	2,430	542
Log-Likelihood: 1.0e+06*	-2.6534	-0.74462	-1.911	-4.2768	-0.48718	-3.7897
Number Iterations	13	16	12	13	12	12

Note: Point estimates correspond to parameter values in log-log link function. t-statistic in parenthesis. All specifications control for locality, year and age-specific fixed effects

Table 5: Main results by gender or age.

	Gender			Age		
	>2.5k	Semi-Urban	>15k	>2.5k	Semi-Urban	>15k
Seguro Popular Male	-0.2246***	-0.1469***	-0.2609***			
	(-21.78)	(-9.55)	(-17.38)			
Seguro Popular Female	-0.0489***	-0.0322**	-0.0691***			
	(-4.71)	(-2.05)	(-4.64)			
Seguro Popular 60-99				-0.1937***	-0.1413***	-0.233***
				(-16.63)	(-8.25)	(-14.00)
Seguro Popular 20-59				-0.0497***	0.0054	-0.0666***
				(-3.45)	(0.24)	(-3.48)
Number Spells	1302512	357772	944740	1302512	357772	944740
Number Localities	3090	2547	543	3090	2547	543
Log-Likelihood: 1.0e+06*	-2.653	-0.74206	-1.9106	-2.6522	-0.74205	-1.9097
Number Iterations	16	12	12	12	12	12

Note: Point estimates correspond to parameter values in log-log link function. t-statistic in parenthesis. All specifications control for locality, year and age-specific fixed effects

Table 6: Main results by gender and age.

	>2.5k	Semi-Urban	>15k
Seguro Popular Male 60-99	-0.3228***	-0.2331***	-0.376***
	(-27.46)	(-13.14)	(-22.74)
Seguro Popular Female 60-99	-0.0163	0.0023	-0.0342**
	(-1.35)	(0.13)	(-2.03)
Seguro Popular Male 20-59	-0.1457***	-0.0804***	-0.1762***
	(-10.42)	(-3.63)	(-9.37)
Seguro Popular Female 20-59	-0.0201	0.0011	-0.0346*
	(-1.35)	(0.05)	(-1.76)
Number Spells	1,302,512	357,772	944,740
Number Localities	3,090	2,547	543
Log-Likelihood: 1.0e+06*	-2.6518	-0.74203	-1.9093
Number Iterations	11	16	13

Note: Point estimates correspond to parameter values in log-log link function. t-statistic in parenthesis. All specifications control for locality, year and age-specific fixed effects

Table 7: Hazard rates of death within the next year, based on results in table 6.

	Localities >2,500			Semi-Urban			Localities >15,000		
	No SP	SP	Diff.	No SP	SP	Diff.	No SP	SP	Diff.
70 Year Old Male	1.59%	1.16%	-0.44%	1.51%	1.20%	-0.31%	1.74%	1.20%	-0.54%
70 Year Old Female	1.75%	1.72%	-0.03%	1.56%	1.56%	0.00%	1.95%	1.89%	-0.07%
40 Year Old Male	0.71%	0.62%	-0.10%	0.78%	0.72%	-0.06%	0.74%	0.62%	-0.12%
40 Year Old Female	1.22%	1.19%	-0.02%	1.23%	1.23%	0.00%	1.30%	1.26%	-0.04%

Note: Hazard rates (yearly probability of death at each age) based on model with age (under 60/ 60 years of age or older) and gender interaction term.

Table 8: Main results interacted with 2002 Health Ministry and Non-Health Ministry staff at public institutions.

	Baseline			Health Ministry			Non-Health Ministry		
	>2.5k	Semi-Urban	>15k	>2.5k	Semi-Urban	>15k	>2.5k	Semi-Urban	>15k
Seguro Popular	-0.124***	-0.0889***	-0.1437***	-0.1245***	-0.0877***	-0.1452***	-0.1013***	-0.0893***	-0.1281***
	(-11.96)	(-5.94)	(-9.42)	(-11.95)	(-5.81)	(-9.45)	(-9.68)	(-5.98)	(-8.35)
Seguro Popular*Health Staff				-5.2454	7.832	-10.3136	-70.7963***	32.2995**	-59.3317***
				(-0.54)	(0.52)	(-0.82)	(-14.28)	(2.39)	(-10.31)
Number Spells	1,082,180	322,111	760,069	1,082,180	322,111	760,069	1,082,180	322,111	760,069
Number Localities	2,863	2,374	489	2,863	2,374	489	2,863	2,374	489
Log-Likelihood: 1.0e+06*	-2.1980	-0.66837	-1.5293	-2.198	-0.66969	-1.5293	-2.198	-6.6836	-1.5293
Number Iterations	13	13	12	13	13	12	13	13	13

Note: Point estimates correspond to parameter values in log-log link function. t-statistic in parenthesis. All specifications control for locality, year and age-specific fixed effects

Table 9: Results for each gender interacted with 2002 Health Ministry and Non-Health Ministry staff at public institutions.

	Baseline			Health Ministry			Non-Health Ministry		
	>2.5k	Semi-Urban	>15k	>2.5k	Semi-Urban	>15k	>2.5k	Semi-Urban	>15k
Seguro Popular Male	-0.2047***	-0.1396***	-0.2362***	-0.205***	-0.1404***	-0.2377***	-0.177***	-0.1402***	-0.2166***
	(-18.63)	(-8.61)	(-14.81)	(-18.57)	(-8.61)	(-14.79)	(-15.79)	(-8.60)	(-13.46)
Seguro Popular Female	-0.0307***	-0.0318*	-0.0357**	-0.0315***	-0.0296*	-0.0378**	-0.0161	-0.0321*	-0.0257
	(-2.78)	(-1.92)	(-2.26)	(-2.84)	(-1.77)	(-2.37)	(-1.42)	(-1.92)	(-1.61)
Seguro Popular*Health Staff Male				1.0854	10.6099	-5.5487	-80.3225***	29.7051	-70.2303***
				(0.09)	(0.54)	(-0.35)	(-12.81)	(1.64)	(-9.67)
Seguro Popular*Health Staff Female				-13.4117	1.7642	-20.7381	-52.8925***	38.081*	-40.9728***
				(-0.95)	(0.08)	(-1.13)	(-7.26)	(1.85)	(-4.81)
Number Spells	1,082,180	322,111	760,069	1,082,180	322,111	760,069	1,082,180	322,111	760,069
Number Localities	2,863	2,374	489	2,863	2,374	489	2,863	2,374	489
Log-Likelihood: 1.0e+06*	-2.1976	-0.66833	-1.5289	-2.1976	-6.6834	-1.5289	-2.1977	-0.66832	-1.5289
Number Iterations	13	13	14	13	14	13	13	13	15

Note: Point estimates correspond to parameter values in log-log link function. t-statistic in parenthesis. All specifications control for locality, year and age-specific fixed effects

References

Azuara, O. & Marinescu, I. (2010), Informality and the expansion of social protection programs. evidence from mexico. unpublished working paper.

Barros, A. J., Santos, I. S. & Bertoldi, A. D. (2008), 'Can mothers rely on the brazilian health system for their deliveries? an assessment of use of the public system and out-of-pocket expenditure in the 2004 pelotas birth cohort study, brazil', *BMC Health Services Research* **8**(57).

Barros, R. (2008), Wealthier but not much healthier: Effects of a health insurance program for the poor in mexico. unpublished working paper.

Bitrán, R., Muñoz, R. & Prieto, L. (2010), Health insurance and access to health services, health services use, and health status in peru, *in* M.-L. Escober, C. C. Griffin & P. R. Shaw, eds, 'The Impact of Health Insurance in Low- and Middle-Income Countries', Brookings Institution Press, Washington, DC, pp. 106–21.

Bleich, S. N., Cutler, D. M., Adams, A. S., Lozano, R. & Murray, C. J. (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), 875–78.

Bosch, M. & Campos-Vázquez, R. M. (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. & Conover, E. (2013), 'Effects of subsidized health insurance on newborn health in a developing country', *Economic Development and Cultural Change* **61**(3), 633–658.

Camacho, A., Conover, E. & Hoyos, A. (2010), Effects of colombia's social protection system on workers' choice between formal and informal employment. unpublished working paper.

Cercone, J., Etoile, P., Pacheco-Jimenes, J. & Briceno, R. (2010), Impact of health insurance on access, use, and health status in costa rica, *in* M.-L. Escober, C. C. Griffin & P. R. Shaw, eds, 'The Impact of Health Insurance in Low- and Middle-Income Countries', Brookings Institution Press, Washington, DC, pp. 89–105.

Conti, G. & Ginja, R. (2016), Health insurance and child health: Evidence from mexico. IZA Discussion Paper No. 10122.

Conti, G., Ginja, R. & delValle, A. (2017), The effects of social programmes on infant mortality: Evidence from mexico. unpublished working paper.

Conti, G., Ginja, R. & Narita, R. (2016), Non-contributory health insurance and household labor supply: Evidence from mexico. unpublished working paper.

Duval-Hernández, R. & Smith-Ramírez, R. (2011), Informality and seguro popular under segmented labor markets. unpublished working paper.

Fahrmeir, L. & Tutz, G. (2001), *Multivariate Statistical Modelling Based on Generalized Linear Models*, Springer Verlag, New York.

Frenk, J., Gonzalez-Pier, E., Gomez-Dantes, O., Lezana, M. A. & Knaul, F. M. (2006), 'Comprehensive reform to improve health system performance in mexico', *The Lancet* **368**, 1524–34.

Gakidou, E., Lozano, R., Gonzalez-Pier, E., Abbott-Klafter, J., Barofsky, J. T., Bryson-Cahn, C., Feehan, D. M., Lee, D. K., Hernandez-Llamas, H. & Murray, C. J. (2006), 'Assessing the effect of the 2001-06 mexican health reform: An interim report card', *The Lancet* **368**, 1920–35.

Giedion, U. & Díaz, B. Y. (2010), A review of the evidence, *in* M.-L. Escober, C. C. Griffin & P. R. Shaw, eds, 'The Impact of Health Insurance in Low- and Middle-Income Countries', Brookings Institution Press, Washington, DC, pp. 13–32.

Giedion, U., Díaz, B. Y., Alfonso, E. A. & Savedoff, W. D. (2009), The impact of subsidized health insurance on health status and on access to and use of health services, *in* A. L. Glassman, M.-L. Escobar, A. Giuffrida & U. Giedion, eds, 'From Few to Many: Ten Years of Health Insurance Expansion in Colombia', Inter-American Development Bank and The Brookings Institution, Washington, DC, pp. 47–74.

Gómez-Morín-Escalante, L. (2015), 'Las muertes invisibles en méxico', Animal Politico. http://www.animalpolitico.com/blogueros-el-blog-de-mexico-evalua/2015/01/09/las-muertes-invisibles-en-mexico/.

Grogger, J., Arnold, T., León, A. S., Ome, A. & Triyana, M. (2010), Identificación y análisis de los efectos del seguro popular en el gasto en salud de los afiliados, Technical report, Centro de Investigación y Docencia Económicas. Evaluation commissioned by the Comisión Nacional de Protección Social en Salud (CNPPS).

Gross, S. T. & Catherine, H.-C. (1992), 'Regression models for truncated survival data', *Scandinavian Journal of Statistics* **19**(3), 193–213.

King, G., Gakidou, E., Imai, K., Lakin, J., Moore, R. T., Nall, C., Ravishankar, N., Vargas, M., Tellez-Rojo, M. M., Hernandez-Avila, J. E., Hernandez-Avila, M. & Hernandez-Llamas, H. (2009), 'Public policy for the poor? a randomized assessment of the mexican universal health insurance programme', *The Lancet* **373**, 1447–54.

Knaul, F. M., Arreola-Ornelas, H., Méndez-Carniado, O., Bryson-Cahn, C., Barofsky, J., Maguire, R., Miranda, M. & Sesma, S. (2006), 'Evidence is good for your health system: Policy reform to remedy catastrophic and impoverishing health spending in mexico', *The Lancet* **368**, 1828–41.

Knox, M. (2008), Health insurance for all: An evaluation of mexico's seguro popular program. Unpublished working paper.

Lakin, J. M. (2010), 'The end of insurance? mexico's seguro popular, 2001-2007', *Journal of Health Politics, Policy and Law* **35**(3).

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

Martínez, G. & Aguilera, N. (2010), The value of health and longevity in mexico. unpublished working paper.

Pfutze, T. (2014), 'The effects of mexico?s seguro popular health insurance on infant mortality: An estimation with selection on the outcome variable', *World Development* **59**, 475–86.

Pfutze, T. (2015), 'Does access to health insurance reduce the risk of miscarriages? evidence from mexico?s seguro popular', *Latin American Economic Review* **24**(8).

Ruvalcaba, L. N. & Parker, S. (2010), Identificación y análisis de los efectos en las condiciones de salud de los afiliados al seguro popular, Technical report, Centro de Investigación y Docencia Económicas. Evaluation commissioned by the Comisión Nacional de Protección Social en Salud (CNPPS).

Ruvalcaba, L. & Vargas, M. (2010), Evolución del sistema de protección social en salud, Technical report, Centro de Investigación y Docencia

Económicas. Evaluation commissioned by the Comisión Nacional de Protección Social en Salud (CNPPS).

Scheike, T. H. & Keiding, N. (2006), 'Design and analysis of time-to-pregnancy', *Statistical Methods in Medical Research* **15**, 127–140.

Scott, J. & Aguilera, N. (2010), Análisis de la factibilidad, alternativas de diseño y consecuencias distributivas de la construcción de un sistema integrado de salud en méxico, Technical report, Centro de Investigación y Docencia Económicas. Evaluation commissioned by the Comisión Nacional de Protección Social en Salud (CNPPS).

Victoria, C. G., Aquino, E. M., doCarmo Leal, M., Monteiro, C. A., Barros, F. C. & Szwarcwald, C. L. (2011), 'Maternal and child health in brazil: Progress and challenges', *The Lancet* **377**, 1863–76.

Victoria, C. G. & Peters, D. H. (2009), 'Seguro popular in mexico: Is premature evaluation healthy?', *The Lancet* **373**, 1404–05.

Appendices

A Implementation

The estimation of the parameter vector β in the linear form presents the challenge that the maximization of the likelihood function resulting from the product of expression 3 over all observations gives rise to unnecessarily complicated gradients and hessian matrix, generalized linear models provides a more tractable framework. First note that for each time-specific observation, following expression 2 above, the likelihood can be written as:

$$L_{i,t}(\beta) = \left[\frac{exp(-F_i(t-1)) - exp(-F_i(t))}{exp(-F_i(t-1)) - exp(-F_i(S_i))}\right]^{y_{i,t}} \left[\frac{exp(-F_i(t)) - exp(-F_i(S_i))}{exp(-F_i(t-1)) - exp(-F_i(S_i))}\right]^{1-y_{i,t}}$$

taking logs and rearranging:

$$l_{i,t}(\beta) = y_{i,t} log\left[\frac{exp(-F_i(t-1)) - exp(-F_i(t))}{exp(-F_i(t)) - exp(-F_i(S_i))}\right] + log\left[\frac{exp(-F_i(t)) - exp(-F_i(S_i))}{exp(-F_i(t-1)) - exp(-F_i(S_i))}\right] \tag{1}$$

This can be put into the general form for c.d.f.s in the exponential family by defining:

$$\theta_{i,t} = log\left(\frac{exp(-F_i(t-1)) - exp(-F_i(t))}{exp(-F_i(t)) - exp(-F_i(S_i))}\right) = log\left(\frac{1 - e^{-e^{x_{i,t}\beta}}}{e^{-e^{x_{i,t}\beta}} - e^{-\sum_{j=t}^{S} e^{x_{i,j}\beta}}}\right)$$

$$b(\theta_{i,t}) = -log\left(\frac{1}{1+e^{\theta_{i,t}}}\right) = -log\left(\frac{e^{-e^{x_{i,t}\beta}} - e^{-\sum_{j=t}^{S} e^{x_{i,j}\beta}}}{1 - e^{-\sum_{j=t}^{S} e^{x_{i,j}\beta}}}\right)$$

Taking the first derivatives w.r.t θ:

$$b'(\theta_{i,t}) = \mu_{i,t} = \frac{e^{\theta_{i,t}}}{1+e^{\theta_{i,t}}} = \frac{exp(-F_i(t-1)) - exp(-F_i(t))}{exp(-F_i(t-1)) - exp(-F_i(S_i))} \qquad (2)$$

Following the (Fahrmeir & Tutz 2001) the likelihood function in equation 1 can now be written as:

$$l_{i,t}(\beta) = y_{i,t}[\theta_{i,t}h(X\beta)] - b(\theta_{i,t}(h(X\beta))),$$

where $h(X\beta)$ is defined by $\mu_{i,t} = h(X\beta)$ in expression 2. The function $\theta_{i,t}(h(.))$ can hence be derived as:

$$\theta_{i,t}(h(.)) = log\left(\frac{h(X\beta)}{1-h(X\beta)}\right)$$

It follows that for a particular parameter β_k, the first and second derivates are:

$$
\begin{aligned}
\frac{\partial l_{i,t}(\beta)}{\beta_k} =& y_{i,t}\frac{d\theta_{i,t}}{dh}\frac{\partial h}{\partial \beta_k} - \frac{db}{d\theta_{i,t}}\frac{d\theta_{i,t}}{dh}\frac{\partial h}{\partial \beta_k}, \\
=& (y_{i,t}-\mu_{i,t})\frac{d\theta_{i,t}}{dh}\frac{\partial h}{\partial \beta_k} \\
\frac{\partial^2 l_{i,t}(\beta)}{\beta_k\beta_l} =& -\frac{\partial h}{\partial \beta_l}\frac{d\theta_{i,t}}{dh}\frac{\partial h}{\partial \beta_k} + (y_{i,t}-\mu_{i,t})\left(-\frac{1}{(\mu_{i,t}(1-\mu_{i,t}))^2}\right)\left(\frac{\partial h}{\partial \beta_l}(1-\mu_{i,t}) - \mu_{i,t}\frac{\partial h}{\partial \beta_l}\right)\frac{\partial h}{\partial \beta_k} \\
&+ (y_{i,t}-\mu_{i,t})\frac{d\theta_{i,t}}{dh}\left(\frac{\partial^2 h}{\partial \beta_k \partial \beta_l}\right) \\
=& \frac{1}{\mu_{i,t}(1-\mu_{i,t})}\left[\left((y_{i,t}-\mu_{i,t})\frac{2\mu_{i,t}-1}{\mu_{i,t}(1-\mu_{i,t})} - 1\right)\frac{\partial h}{\partial \beta_k}\frac{\partial h}{\partial \beta_l} + (y_{i,t}-\mu_{i,t})\frac{\partial^2 h}{\partial \beta_k \partial \beta_l}\right]
\end{aligned} \tag{3}
$$

The precise form of the different elements of the expressions will be shown below.

B Detailed Gradients and Hessians for Newton-Raphson Method

Here, the different elements of the first and second order conditions in equation 2 will be derived. Noting that h(.) can be simplified to $h(x_{it}\beta) = \frac{1-e^{-e^{x_{i,t}\beta}}}{1-e^{-\sum_{j=t}^{S} e^{x_{i,j}\beta}}}$ the elements of the first derivative can be expressed as:

$$\frac{d\theta_{i,t}}{dh} = \frac{1}{\mu_{i,t}(1-\mu_{i,t})}$$

$$\frac{\partial h}{\partial \beta_k} = \frac{e^{-e^{x_{i,t}\beta}} e^{x_{i,t}\beta} x_{k,i,t} \left(1 - e^{-\sum_{j=t}^{S_i} g(\delta,a_j)e^{x_{i,j}\beta}}\right) - \left(1 - e^{-e^{x_{i,t}\beta}}\right) e^{-\sum_{j=t}^{S_i} g(\delta,a_j)e^{x_{i,j}\beta}} \sum_{j=t}^{S_i} g(\delta,a_j) e^{x_{i,j}\beta} x_{k,i,j}}{\left(1 - e^{-\sum_{j=t}^{S_i} g(\delta,a_j)e^{x_{i,j}\beta}}\right)^2}$$

$$= \frac{\left(e^{-e^{x_{i,t}\beta}} - e^{-\sum_{j=t}^{S_i} g(\delta,a_j)e^{x_{i,j}\beta}}\right) e^{x_{i,t}\beta} x_{k,i,t} + e^{-\sum_{j=t}^{S_i} g(\delta,a_j)e^{x_{i,j}\beta}} \left(e^{-e^{x_{i,t}\beta}} - 1\right) \sum_{j=t+1}^{S_i} e^{x_{i,j}\beta} x_{k,i,j}}{\left(1 - e^{-\sum_{j=t}^{S_i} g(\delta,a_j)e^{x_{i,j}\beta}}\right)^2}$$

Moving forward, the second-order derivatives for the likelihood function are:

$$
\begin{aligned}
&\frac{\partial^2 h}{\partial \beta_k \partial \beta_l} = \\
&\frac{\left[\left(-e^{-e^{x_{i,t}\beta}} e^{x_{i,t}\beta} x_{l,i,t} + e^{-\sum_{j=t}^{S_i} g(\delta,a_j) e^{x_{i,j}\beta}} \sum_{j=t}^{S_i} g(\delta,a_j) e^{x_{i,j}\beta} x_{l,i,j}\right) e^{x_{i,t}\beta} x_{k,i,t} + \left(e^{-e^{x_{i,t}\beta}} - e^{-\sum_{j=t}^{S_i} g(\delta,a_j) e^{x_{i,j}\beta}}\right) e^{x_{i,t}\beta} x_{k,i,t} x_{l,i,t}\right] \left(1 - e^{-\sum_{j=t}^{S_i} g(\delta,a_j) e^{x_{i,j}\beta}}\right)^2}{\left(1 - e^{-\sum_{j=t}^{S_i} g(\delta,a_j) e^{x_{i,j}\beta}}\right)^4} \\
&+ \frac{\left[-e^{-\sum_{j=t}^{S_i} g(\delta,a_j) e^{x_{i,j}\beta}} \sum_{j=t}^{S_i} g(\delta,a_j) e^{x_{i,j}\beta} x_{l,i,j} \left(e^{-e^{x_{i,t}\beta}} - 1\right) \sum_{j=t+1}^{S_i} e^{x_{i,j}\beta} x_{k,i,j} - e^{-\sum_{j=t}^{S_i} g(\delta,a_j) e^{x_{i,j}\beta}} e^{-e^{x_{i,t}\beta}} e^{x_{i,t}\beta} x_{l,i,t} \sum_{j=t+1}^{S_i} e^{x_{i,j}\beta} x_{k,i,j}\right] \left(1 - e^{-\sum_{j=t}^{S_i} g(\delta,a_j) e^{x_{i,j}\beta}}\right)^2}{\left(1 - e^{-\sum_{j=t}^{S_i} g(\delta,a_j) e^{x_{i,j}\beta}}\right)^4} \\
&+ \frac{e^{-\sum_{j=t}^{S_i} g(\delta,a_j) e^{x_{i,j}\beta}} \left(e^{-e^{x_{i,t}\beta}} - 1\right) \sum_{j=t+1}^{S_i} e^{x_{i,j}\beta} x_{l,i,j} x_{k,i,j} \left(1 - e^{-\sum_{j=t}^{S_i} g(\delta,a_j) e^{x_{i,j}\beta}}\right)^2}{\left(1 - e^{-\sum_{j=t}^{S_i} g(\delta,a_j) e^{x_{i,j}\beta}}\right)^4} \\
&- \frac{\left[\left(e^{-e^{x_{i,t}\beta}} - e^{-\sum_{j=t}^{S_i} g(\delta,a_j) e^{x_{i,j}\beta}}\right) e^{x_{i,t}\beta} x_{k,i,t} + e^{-\sum_{j=t}^{S_i} g(\delta,a_j) e^{x_{i,j}\beta}} \left(e^{-e^{x_{i,t}\beta}} - 1\right) \sum_{j=t+1}^{S_i} e^{x_{i,j}\beta} x_{k,i,j}\right] 2 \left(1 - e^{-\sum_{j=t}^{S_i} g(\delta,a_j) e^{x_{i,j}\beta}}\right) e^{-\sum_{j=t}^{S_i} g(\delta,a_j) e^{x_{i,j}\beta}} \sum_{j=t}^{S_i} g(\delta,a_j) e^{x_{i,j}\beta} x_{l,i,j}}{\left(1 - e^{-\sum_{j=t}^{S_i} g(\delta,a_j) e^{x_{i,j}\beta}}\right)^4}
\end{aligned}
$$

C Calculating Gradients and Hessians for Wald test

From 1, the log-likelihood contribution of each, year specific observation is, depending on the outcome:

$$\mathbf{y=1}: l = log[exp(-F(t-1)) - exp(-F(t))] - log[exp(-F(t-1)) - exp(-F(S))]$$

$$\mathbf{y=0}: l = log[exp(-F(t)) - exp(-F(S))] - log[exp(-F(t-1)) - exp(-F(S))]$$

Suppressing spell specific i subscripts, let $E(t) = exp(-F(t))$ and $e(t) = \sum_{j=E}^{t} e^{X_j\beta} X_{j,k}$. Then:

Gradient:

$$\begin{aligned}
&\frac{\partial \mathbf{l}}{\partial \beta_{\mathbf{k}}}: \\
\mathbf{y=1}: \; &= \frac{-E(t-1)e_k(t-1) + E(t)e_k(t)}{E(t-1) - E(t)} - \frac{-E(t-1)e_k(t-1) + E(S)e_k(S)}{E(t-1) - E(S)} \\
&= \frac{E(t)e^{X_t\beta}X_{t,k}}{E(t-1) - E(t)} - \frac{E(S)\sum_{j=t}^{S} e^{X_j\beta}X_{j,k}}{E(t-1) - E(S)} \\
\mathbf{y=0}: \; &= \frac{-E(t)e_k(t) + E(S)e(s)}{E(t) - E(S)} - \frac{-E(t-1)e_k(t-1) + E(S)e_k(S)}{E(t-1) - E(S)} \\
&= -e^{X_t\beta}X_{t,k} + \frac{E(S)\sum_{j=t+1}^{S} e^{X_j\beta}X_{j,k}}{E(t) - E(S)} - \frac{E(S)\sum_{j=t}^{S} e^{X_j\beta}X_{j,k}}{E(t-1) - E(S)}
\end{aligned}$$

Hessian:

$$\frac{\partial \mathbf{l}}{\partial^{\mathbf{2}} \beta_{\mathbf{k}}}:$$

$$\mathbf{y=1}: \; = \frac{(-E(t)e_k(t)e^{X_t\beta}X_{t,k} + E(t)e^{X_t\beta}X_{t,k}^2)(E(t-1)-E(t)) - E(t)e^{X_t\beta}X_{t,k}(-e_k(t-1)(E(t-1)-E(t)) + E(t)e^{X_t\beta}X_{t,k})}{(E(t-1)-E(t))^2}$$

$$- \frac{(-E(S)e_k(S)\sum_{j=t}^{S} e^{X_j\beta}X_{j,k} + E(S)\sum_{j=t}^{S} e^{X_j\beta}X_{j,k}^2)(E(t-1)-E(S))}{(E(t-1)-E(S))^2}$$

$$+ \frac{E(S)\sum_{j=t}^{S} e^{X_j\beta}X_{j,k}(-e_k(t-1)(E(t-1)-E(S)) + E(S)\sum_{j=t}^{S} e^{X_j\beta}X_{j,k})}{(E(t-1)-E(S))^2}$$

$$= \frac{\mathbf{E(t-1)E(t)(e^{X_t\beta}X_{t,k}^2 - (e^{X_t\beta}X_{t,k})^2) - E(t)^2e^{X_t\beta}X_{t,k}^2}}{\mathbf{(E(t-1)-E(t))^2}}$$

$$+ \frac{\mathbf{E(t-1)E(S)((\sum_{j=t}^{S} e^{X_j\beta}X_{j,k})^2 - \sum_{j=t}^{S} e^{X_j\beta}X_{j,k}^2) + E(S)^2\sum_{j=t}^{S} e^{X_j\beta}X_{j,k}^2}}{\mathbf{(E(t-1)-E(S))^2}}$$

$$\mathbf{y=0}: \; = -e^{X_t\beta}X_{t,k}^2 + \frac{(-E(S)e_k(S)\sum_{j=t+1}^{S} e^{X_j\beta}X_{j,k} + E(S)\sum_{j=t+1}^{S} e^{X_j\beta}X_{j,k}^2)(E(t)-E(S))}{(E(t)-E(S))^2}$$

$$- \frac{E(S)\sum_{j=t+1}^{S} e^{X_j\beta}X_{j,k}(-e_k(t)(E(t)-E(S)) + E(S)\sum_{j=t+1}^{S} e^{X_j\beta}X_{j,k})}{(E(t)-E(S))^2}$$

$$- \frac{(-E(S)e_k(S)\sum_{j=t}^{S} e^{X_j\beta}X_{j,k} + E(S)\sum_{j=t}^{S} e^{X_j\beta}X_{j,k}^2)(E(t-1)-E(S))}{(E(t-1)-E(S))^2}$$

$$+ \frac{E(S)\sum_{j=t}^{S} e^{X_j\beta}X_{j,k}(-e_k(t-1)(E(t-1)-E(S)) + E(S)\sum_{j=t}^{S} e^{X_j\beta}X_{j,k})}{(E(t-1)-E(S))^2}$$

$$= -\mathbf{e^{X_t\beta}X_{t,k}^2} + \frac{\mathbf{E(t)E(S)(\sum_{j=t+1}^{S} e^{X_j\beta}X_{j,k}^2 - (\sum_{j=t+1}^{S} e^{X_j\beta}X_{j,k})^2) - E(S)^2\sum_{j=t+1}^{S} e^{X_j\beta}X_{j,k}^2}}{\mathbf{(E(t)-E(S))^2}}$$

$$+ \frac{\mathbf{E(t-1)E(S)((\sum_{j=t}^{S} e^{X_j\beta}X_{j,k})^2 - \sum_{j=t}^{S} e^{X_j\beta}X_{j,k}^2) + E(S)^2\sum_{j=t}^{S} e^{X_j\beta}X_{j,k}^2}}{\mathbf{(E(t-1)-E(S))^2}}$$

$$\frac{\partial \mathbf{l}}{\partial \beta_{\mathbf{k}} \partial \beta_{\mathbf{l}}}:$$

$$\begin{aligned}
\mathbf{y=1}:\ &= \frac{(-E(t)e_l(t)e^{X_t\beta}X_{t,k} + E(t)e^{X_t\beta}X_{t,k}X_{t,l})(E(t-1)-E(t)) - E(t)e^{X_t\beta}X_{t,k}(-e_l(t-1)(E(t-1)-E(t)) + E(t)e^{X_t\beta}X_{t,l})}{(E(t-1)-E(t))^2} \\
&- \frac{(-E(S)e_l(S)\sum_{j=t}^{S} e^{X_j\beta}X_{j,k} + E(S)\sum_{j=t}^{S} e^{X_j\beta}X_{j,k}X_{j,l})(E(t-1)-E(S))}{(E(t-1)-E(S))^2} \\
&+ \frac{E(S)\sum_{j=t}^{S} e^{X_j\beta}X_{j,k}(-e_l(t-1)(E(t-1)-E(S)) + E(S)\sum_{j=t}^{S} e^{X_j\beta}X_{j,l})}{(E(t-1)-E(S))^2} \\
&= \frac{\mathbf{E(t-1)E(t)(e^{X_t\beta}X_{t,k}X_{t,l} - e^{2X_t\beta}X_{t,k}X_{t,l}) - E(t)^2e^{X_t\beta}X_{t,k}X_{t,l}}}{\mathbf{(E(t-1)-E(t))^2}} \\
&+ \frac{\mathbf{E(t-1)E(S)(\sum_{j=t}^{S} e^{X_j\beta}X_{j,k}\sum_{j=t}^{S} e^{X_j\beta}X_{j,l} - \sum_{j=t}^{S} e^{X_j\beta}X_{j,k}X_{j,l}) + E(S)^2\sum_{j=t}^{S} e^{X_j\beta}X_{j,k}X_{j,l}}}{\mathbf{(E(t-1)-E(S))^2}} \\
\mathbf{y=0}:\ &= -e^{X_t\beta}X_{t,k}X_{t,l} + \frac{(-E(S)e_l(S)\sum_{j=t+1}^{S} e^{X_j\beta}X_{j,k} + E(S)\sum_{j=t+1}^{S} e^{X_j\beta}X_{j,k}X_{j,l})(E(t)-E(S))}{(E(t)-E(S))^2} \\
&- \frac{E(S)\sum_{j=t+1}^{S} e^{X_j\beta}X_{j,k}(-e_l(t)(E(t)-E(S)) + E(S)\sum_{j=t+1}^{S} e^{X_j\beta}X_{j,l})}{(E(t)-E(S))^2} \\
&- \frac{(-E(S)e_l(S)\sum_{j=t}^{S} e^{X_j\beta}X_{j,k} + E(S)\sum_{j=t}^{S} e^{X_j\beta}X_{j,k}X_{j,l})(E(t-1)-E(S))}{(E(t-1)-E(S))^2} \\
&+ \frac{E(S)\sum_{j=t}^{S} e^{X_j\beta}X_{j,k}(-e_l(t-1)(E(t-1)-E(S)) + E(S)\sum_{j=t}^{S} e^{X_j\beta}X_{j,l})}{(E(t-1)-E(S))^2} \\
&= \mathbf{-e^{X_t\beta}X_{t,k}X_{t,l}} + \frac{\mathbf{E(t)E(S)(\sum_{j=t+1}^{S} e^{X_j\beta}X_{j,k}X_{j,l} - \sum_{j=t+1}^{S} e^{X_j\beta}X_{j,k}\sum_{j=t+1}^{S} e^{X_j\beta}X_{j,l}) - E(S)^2\sum_{j=t+1}^{S} e^{X_j\beta}X_{j,k}X_{j,l}}}{\mathbf{(E(t)-E(S))^2}} \\
&+ \frac{\mathbf{E(t-1)E(S)(\sum_{j=t}^{S} e^{X_j\beta}X_{j,k}\sum_{j=t}^{S} e^{X_j\beta}X_{j,l} - \sum_{j=t}^{S} e^{X_j\beta}X_{j,k}X_{j,l}) + E(S)^2\sum_{j=t}^{S} e^{X_j\beta}X_{j,k}X_{j,l}}}{\mathbf{(E(t-1)-E(S))^2}}
\end{aligned}$$