Rendering Citizens Legible?

Birth Certificates as a Measure of State Capacity in Mexico

Imke Harbers
Department of Political Science
University of Amsterdam

i.harbers@uva.nl

Version: December 4, 2017

Paper prepared for a Meeting of the Research Seminar at Trinity College, Dublin

This paper constitutes a first attempt to write up and organize the results of the analysis of birth records in Mexico. I greatly appreciate feedback and comments. Please ask for an updated version of the paper before citing or circulating.

Introduction

Contemporary states are engaged in extensive efforts to collect information about populations and territory. Accurate, complete and up-to-date information is a precondition for effective governance (Scott 1998). The state's ability "to implement logistically political decisions throughout the realm" (Mann 1988: 5) depends crucially on information. States therefore attempt "with varying success to create a terrain and a population with precisely those standardized characteristics that will be easiest to monitor, count, assess, and manage" (Scott 1998: 81). While information about the terrain is generally recorded in cadastral records and maps, data about populations is collected periodically by means of a census and continuously with vital registries. Even though all states require information to govern, the extent to which states actually possess accurate information varies widely, not only across countries, but also within them.

Empirical research on the "informational foundations of state capacity" (Lee and Zhang 2016) has flourished in recent years. The frequency of census taking (Soifer 2013), census accuracy (Lee and Zhang 2016), the quality and coverage of cadasters (D'Arcy and Nistotskaya 2017) and the universality of birth registration (Hunter and Brill 2016) have been found to vary across countries, even among those with similar levels of economic development. So far, this research has tended to focus on the coverage and content of public records at the national level. Conceptually, such aggregate indicators measure deviations from the ideal of accurate, complete and up-to-date information. This paper examines the micro-foundations of aggregate indicators, and unpacks how state information varies within the territory and across societal groups. Focusing on Mexico, it examines one key aspect of state information about society: the ability to accurately identify individual citizens. This ability depends on the quality of vital records, specifically birth registration.

Birth certificates legally recognize the existence of an individual and establish and protect individual identity. They not only provide the state with information about demographic trends in society but, in contrast to the census, they also confer rights to their holders. The inability of the state to identify individual citizens has substantial economic, political, and social implications. For one, the capacity to enforce contracts and uphold property rights depends on the state's ability to ascertain individual identities. Democratic governance is only possible where election officials can effectively monitor whether those eligible were allowed to vote, and whether nobody voted more than once. This requires a complete and up-to-date registry of the population's nationality and age. Where voter rolls are of poor quality, the electoral process is vulnerable to fraud or allegations of fraud. With regard to social policy, the capacity of the state to implement need-based programs (rather than rely on clientelism or patronage) depends on its ability to accurately target benefits, and to monitor the distribution of funds (Hunter and Brill 2016). Where the state is unable to accurately identify populations in need, this undermines its capacity to engage in redistribution and poverty alleviation. In the absence of accurate information, all state interventions in society are "necessarily crude" (Scott 1998: 77). Moreover, where the state is unable to credibly commit to redistribution, this undermines support for democracy among marginalized populations who have little to gain from participating in the democratic process (Soifer 2013). A special issue of 'The Lancet' devoted to vital registries goes so far as to argue that vital registration is not only a precondition for effective public good provision, it constitutes a public good in and of itself (Setel et al 2007). Yet, while there is a strong consensus about its importance, fairly little is known about the factors

that determine whether citizens have access to timely birth registration, and why the quality of records varies across societal groups.

Even though Mexico, along with many other Latin American countries, has increased access to birth registration over the past decade, as recently as 2009 the director of Mexico's civil registry estimated that 1 in 5 Mexicans is not registered and thus effectively "undocumented" in his or her own country (Valdés 2011: 11). This speaks to a marked unevenness in the "intensity of citizenship" across society (O'Donnell 1993). Some citizens interact regularly with the state, receive services to which they are entitled and thus possess documents to prove their identity. Others, however, have only limited contact with state agencies and documents are of limited value for them.

The paper draws on an original dataset of roughly 80 million records for births registered between 1985 and 2005. Birth certificates contain basic demographic information about the date of birth, date of registration, gender as well as geographical identifiers for municipalities, delegations or (in the case of Oaxaca) districts. In its current form, this paper is still rather descriptive. Cleaning the information in the certificates to ensure that codes are consistent across time and space has been very time-intensive, and work is still ongoing for some of the demographic variables. To analyze whether the information in the birth records can reasonably be used by the state to pinpoint individuals, the paper's methodological approach is twofold. First, information in the birth records is cross-validated with other societal information collected by the state, specifically census data and population projections. Second, demographic patterns in the data are checked for consistency with demographic trends. Both approaches indicate significant problems with the data, and cast doubt on the Mexican state's ability to identify individual citizens.

The paper shows that in the absence of universal and timely registration, the most vulnerable citizens are likely to be remain "unseen" by the state, which in turn limits the state's ability to effectively implement policies for those in need. Moreover, the paper shows that increasing access to birth registration does not necessarily increase the state's ability to accurately identify individual citizens, especially among vulnerable populations. While access to birth certificates has increased and more people now have records, these records are less likely to be accurate (content errors), and they are less likely to uniquely identify individual citizens (coverage errors).

Why do states try to collect data about populations, and when do they succeed?

Data on demographic trends generally come from two main sources: the census and vital registries.[1] A census takes a snapshot of the population at a specific point in time. Vital or civil registration, by contrast, records vital events in the lives of citizens, such as births, marriages or deaths, in real-time (or as close to real-time as possible). Both, in principle, aim for universality, meaning that they intend to cover all people in the population and all parts of the country. A citizen obtains no clear benefit from participating in a census, other than maybe the remote hope that information provided may lead to better services in the future. Early censuses were

[1] States also supplement these sources with surveys among specific sub-groups of the population, but this information is more limited in scope.

therefore often highly contentious and conducted primarily to extract resources from society. Census-taking is a crucial aspect of state development, and the historical literature on state formation is rife with examples of local resistance to this aspect of "state-making" (Soifer 2013).

Vital registries, by contrast, "create and maintain legal documents proving the identity of individuals" (Mahapatra et al 2007: 1653). They not only provide the state with information about society, but also confer rights to their holders. Even though census information may be quite detailed, registration of vital events is the prerequisite for individual legal claims. In modern states "[i]f you wish to have any standing in law, you must have a document that officials accept as evidence of citizenship, be that document a birth certificate, passport, or identity card" (Scott 1998: 83). Birth registration formally recognizes the existence of an individual and establishes and protects individual identity. This, in turn, is the basis for political rights granted to citizens, such as nationality, the right to vote or to obtain a passport, and for protections under the law, such as property rights or access to entitlements.

National level measures of state capacity examine whether state information about terrain and society suffers from systematic blind spots. These blind spots can be territorial. When cadasters cover less than 75% of the territory (D'Arcy and Nistotskaya 2017) or when a national census is limited to urban areas (Soifer 2013), the state lacks accurate information about parts of the country. Blind spots can also be societal and indicate that state information about certain segments of the population is incorrect or incomplete. When census-takers cannot accurately determine the age of respondents (Lee and Zhang 2016), or where birth records are not universal (Hunter and Brill 2016), some societal groups apparently lack official documents. The state then does not possess "a complete and legible list of subjects and taxpayers" (Scott 1998: 69).

Comparative research has shown that comprehensive and accurate state records are the exception, not the rule. Moreover, neither the level of economic development nor technology sufficiently explain the quality of records. Some countries were able to achieve near-universality with ink and paper, while middle income countries like Mexico continue to grapple with poor records despite the advent of computers (Hunter and Brill 2016). In a special issue of *The Lancet* a group of doctors and public health officials make a passionate plea for the improvement of vital registries across the world, and argue that the "establishment of a civil registration system is an act of political will, a demonstration by national authorities of stewardship, and of reciprocal trust in their government by the registered population" (Setel 2017: X). Why have states been unable to muster this political will? And whose information is obscured by blind spots in the state's picture of society?

To better understand why the quality of information varies within countries, it is helpful to acknowledge that state capacity does not emerge organically. Developing and maintaining the infrastructure necessary to record all vital events in the population is costly, since it requires a functioning and minimally competent bureaucracy everywhere. This is one of the reasons why states in the developing world often prefer to rely on census information to gauge demographic trends, or opt for cheaper periodic surveys. While public health officials consider vital registries the gold standard for demographic information, for many states the fixed costs of maintaining the necessary system are simply perceived as too high (Setel et al 2007).

Contrary to public health officials neither states nor citizens consider birth records an end in itself, even though these records potentially offer substantial benefits to both parties. How the costs of collecting information for states interact with the incentives for citizens to provide the information can help account for variation in the quality of records.

Birth registration sheds light on the interaction between state and society. From the state's perspective, given limited resources, it makes sense to collect only those facts that are directly relevant to state projects and policies. Since Prussia was concerned about draft dodging, for example, it recorded the age and gender of citizens, but not their religious denomination (Scott 1998: 52). The contemporary German state, by contrast, collects membership fees for major churches through the so-called "church tax" (*Kirchensteuer*), and therefore needs and records the religious affiliation of tax payers. States are selective in the type of information they record, and in who gets included in the count. Ultimately, information gathering is thus driven by state incentives to record specific types of information.

Birth registration also imposes costs on citizens. Often, registration requires payment of a fee, and parents (or citizens seeking documents) have to make a trip to the registrar's office. For poor parents, and those in remote parts of the country, the monetary cost of the fee and travel can already be daunting. This is compounded by the need to take time off from work (often for both parents) and requirements to provide documentation. Less tangible obstacles are often prior experiences with corrupt or unresponsive public officials, which serve as a further disincentive for registration. Populations who, for historical reasons, are distrustful of the state, are less likely to register (Mahapatra 2007). Given these obstacles, it is worth pointing out that the benefits of registration are often not immediately clear, especially for citizens outside the formal economy. Unevenness across societal groups thus results from a situation where the costs of registration are particularly hard to bear for certain societal groups, and their incentives for seeking registration, especially for young children, are low since doing so offers few tangible benefits.

While this overall calculus remains the same, concrete incentives for the state and citizens can shift over time, especially with changing state priorities. With regard to birth registration, Hunter and Brill (2016) argue that the expansion of access to birth records across Latin America is the result of a shift in public policy and changing welfare policies. Whereas welfare spending for the poor had previously occurred primarily via clientelistic or patronage networks, the shift to technocratic entitlements in the form of conditional cash transfers (CCTs) in the late 1990s imposed different information requirements. Where discretionary spending had relied on local knowledge and personal relationships, and was therefore mediated by brokers, the targeting of CCTs required simplified information about populations in need that was 'legible' for technocrats in central government bureaucracies. States began to earnestly invest in expanding access to birth registration, Hunter and Brill argue, only when they reached an administrative barrier to implementing desired social policies. For poor citizens, the implementation of CCT programs offered a tangible benefit that required official documents.

Birth Registration in Mexico

The Mexican state started to wrestle information about populations from the Catholic Church in the 19th century, which guarded this data and the authority to certify vital events jealously. The

1859 *Ley Orgánica del Registro Civil* constituted a first attempt by the state to establish its own civil registry, but coverage and compliance increased only very slowly. Even in 1980, a WHO mission to Mexico was unable to effectively evaluate the quality of vital registration and concluded that the "outputs of the system for the number of births are not reliable because the coverage is unknown". Since there was no hard data on the number of births, it was impossible to assess how many of them were effectively registered and in which time frame.

Even today a UN assessment of census quality puts Mexico in the middle category on a five-point scale, rating the quality of census data as "approximate" based on the Whipple Index for age heaping.[2] This index, which has been put forward by Lee and Zhang (2016) as a measure of state capacity, identifies the extent of age misreporting by examining spikes in the age distribution. These spikes are the result of respondents' or enumerators' propensity to guess round numbers when the exact age is unknown. Birth records are logically prior to heaping in the census. When respondents do not have birth records, and thus no official document with their date of birth, they are less likely to recall their exact age or birthdate when census-takers ask.

For public health officials, the lack of reliable data on birth and deaths in Mexico made it difficult to calculate even the most basic epidemiologic statistics. Qualitative investigations and case studies in specific communities illustrate how inadequate information from vital registries was, especially with regard to infant mortality. In one study, public health officials estimate that twenty percent of deaths before the age of five in poor municipalities had remained unregistered. All children identified by the study whose deaths had not previously been recorded had not been registered at birth, and did not possess birth certificates when they died (Hernández et al 2002: 396). The authors found that traditional midwives – who attended the majority of births in these municipalities – generally did not issue certificates because they did not speak Spanish, were unfamiliar with the form, or unable to read and write. Parents in these municipalities did not ask for documents, because they were unlikely to need a birth certificate prior to enrolling children in school.

While public health officials lamented the dearth of reliable data, politicians and policy scholars tended to shrug and attribute the lack of compliance with registration requirements to insufficient civicmindedness among citizens ("*falta de cultura*").[3] The obstacles to registration for poor and indigenous citizens, and the lack of incentives to seek documentation, received scant attention in these debates.

That the under-registration of births and deaths among Mexico's marginalized citizens not only resulted from bureaucratic ineptness, or the unwillingness of the population to contribute to the crucial public good of accurate vital records, became painfully clear with the 1994 Zapatista uprising. The movement, which mobilized around demands of social justice for indigenous peoples in the impoverished Southern state of Chiapas, claimed that the state had willfully neglected segments of the population. In response to the government's offer to pardon Zapatista troops Subcommandante Marcos, one of the group's spokespersons, responded:

[2] UN Statistics Division, map generated for the workshop "Census Data Evaluation for English-Speaking Countries" based on Data from the Demographic Yearbook (https://unstats.un.org/unsd/demographic/products/dyb/dybcens.htm).

[3] A volume edited by Luz Maria Valdes (2011) to commemorate the 150th anniversary of the civil registry in Mexico outlines many of these debates.

> "…should we ask pardon from the dead, our dead, who died "natural" deaths of "natural causes" like measles, whooping cough, break-bone fever, cholera, typhus, mononucleosis, tetanus, pneumonia, malaria and other lovely gastrointestinal and pulmonary diseases? Our dead, so very dead, so democratically dead from sorrow because no one did anything, because the dead, our dead, went just like that, *with no one keeping count* with no one saying, "Enough!" which would at least have granted some meaning to their deaths, a meaning no one ever sought for them" (emphasis added)

The uprising, and a fear of further unrest, contributed to a shift in social policy, which ultimately cumulated in the creation of Oportunidades, one of the largest conditional cash transfer programs. The program provides better targeted, less partisan, benefits to the poor than previous policies. It is credited not only with reducing infant mortality among vulnerable populations, but also with increasing birth registration (Diaz Cayeros et al 2016; Hunter and Brill 2016).

Evaluating Birth Records in Mexico: Coverage and Content

Demographic data about populations can be affected by errors of *coverage* and of *content*. Coverage errors indicate either the omission of people who should have been included, or the duplication of records. The failure to collect data in parts of the country, or counting some people twice, are coverage errors. Identifying the existence and extent of coverage errors generally requires an alternative source of information that can reasonably be expected to come closer to the real number. Content errors indicate incorrect or incomplete information in the recorded data. Demographers have developed a variety of tools to gauge the extent of content errors. Commonly, they look for patterns in the data that are demographically implausible. Examples of such measures are indices that gauge the prevalence of age heaping, where the normally smooth distribution of age across the population displays sudden spikes. Both types of errors clearly undermine the value of information collected by the state. The following section draws on insights from demography to evaluate the quality of birth records and whether the information can effectively be used by the state to identify individual citizens.

Exploring Coverage: Comparing Aggregate Data

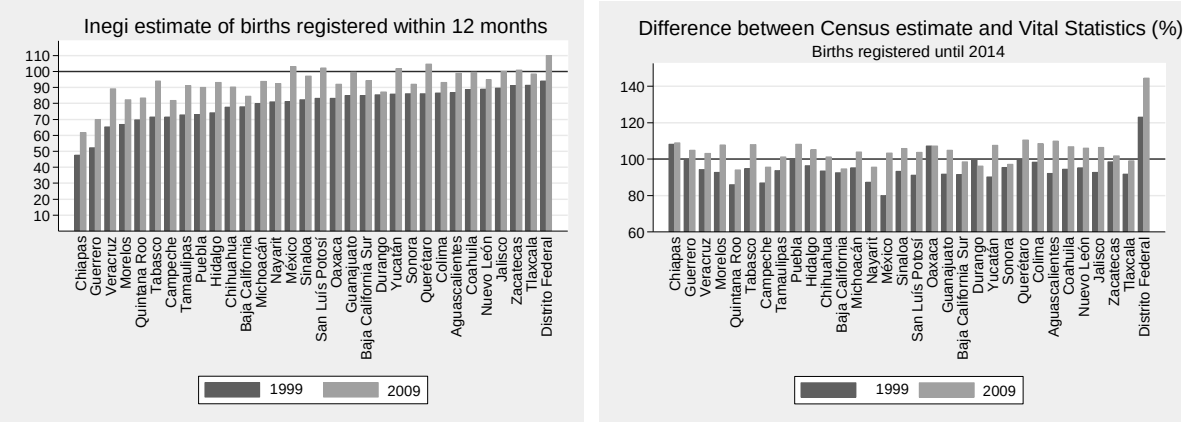
To estimate the coverage of birth records, registered births have to be compared to an alternative data source that better captures the true number of births. In the case of Mexico, two alternatives are available: the national census, which is conducted every ten years, and population projections, calculated by the Consejo Nacional de la Población (CONAPO). Both will be explored in turn.

The Mexican census contains a question asking women with children whether a child was born in the previous year. This is the best available count of the number of births, but it is only available for years preceding a census. The Instituto Nacional de Estadística, Geografía e Informática (INEGI), Mexico's statistical agency, uses data from the census to estimate how many children counted in the census are registered within the first year of their life. This is defined by INEGI (2011) as "timely registration", one of the key indicators for the quality of vital

statistics in the public health literature (Mahapatra 2007).[4] INEGI data show that timely registration has increased significantly between 1999 and 2009, the years prior to the 2000 and 2010 census. In 1999, just 79% of births were registered in the first year. By 2009 that percentage had increased to 93%. These aggregate figures hide significant subnational variation, though. As Figure 1a show, in 1999, around 50% of births in the states of Chiapas and Guerrero were registered within the first year, compared to more than 90% in the Federal District and the states of Tlaxcala and Zacatecas.

Figure 1: Expanding Coverage and Timely Registration

1a) *1b)*



Sources: *Inegi (2011)* & *Estadísticas Vitales*

While all states increased rates of timely registration between 1999 and 2009, the rate of improvement is uneven across states. Also, it is puzzling that some states exceed 100 percent in 2009. This indicates that more children were registered than had been counted in the census. To make sense of this, it is necessary to recognize that INEGI does not individually match babies counted in the census to registered births. The methodology, which is similar to the one employed by Hunter and Brill (2016) in their study of birth registration in Brazil and Bolivia, conceives of each registered birth as progress towards the goal of universality. All registrations are thus regarded as equally valid.

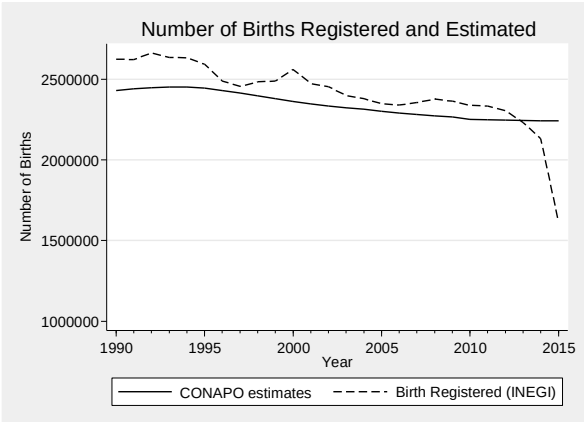
Since aggregate numbers for the country as a whole fall short of 100 percent, it is possible that the excess in some states results from a mismatch between the state of residence, the state of birth and the state of registration. Even though the comparison is based on the state of residence recorded on the birth certificate and state of residence in the census, it is possible that families with children born and counted in one state, Puebla for instance, subsequently register their children in neighboring Estado de México or the Federal District, and misreport their residence. Since urban areas (like the Federal District and surrounding Estado de México) have a greater density of hospitals and other facilities, cross-state mobility could account for the apparent surplus in some states.

[4] The legal framework surrounding birth registration is set in the civil codes of Mexico's 31 states and the Federal District. The time allowed for registration without incurring a late penalty for delayed registration (*registro tardío o extemporáneo*) varies and ranges from as little as 15 days to 365 days. The standard of 12 months employed by INEGI does therefore not correspond to the legal timeframe for "timely registration".

This explanation is undermined by Figure 1b, however. The graph compares the census count to vital statistics, but this time includes not just the first 12 months of a baby's life, but all births reported to have occurred in 1999 and 2009 registered by 2014. The ordering of states is the same as in the previous graph. It is remarkable that Chiapas as well as the Federal District, the top and bottom performer with regard to timely registration, exceed 100 percent in 1999 and 2009. That the predominantly poor and rural states of Chiapas and Oaxaca, and for 2009 also Guerrero, Veracruz and Tabasco, exceed 100 percent rules out internal migration as a plausible explanation. For these states, coverage errors are more plausible as reasons for the surplus. Coverage errors would occur, for instance, if parents lose birth records for their children, and rather than seeking to replace the documents, register the child again. Parents might also not remember that their child had already been registered, since they had never used the child's documents.

Further evidence for the pervasiveness of double registration comes from a comparison of births reported to have occurred between 1990 and 2014 and reported by 2015, and the population projections calculated by CONAPO on the basis of a demographic model. The solid line indicates how many births CONAPO estimates for each year. The dashed line captures births registered for each year by 2014. Two things are noteworthy about the graph. First, the CONAPO line is smooth, which is consistent with demographic patterns. Barring extraordinary circumstances, in any given year the number of births should be roughly the average of those in the previous and following year. The dashed line is much more uneven, and has a fairly sharp peak in 2000. This constitutes evidence for 'heaping', a form of misreporting that results from preferences for round numbers, such as the year 2000, when the exact number is unknown or not regarded as relevant. From a demographic perspective, the solid line is therefore more plausible than the dashed line. Second, for all but the last three years, the number of births registered exceeds the number of births estimated to have occurred by CONAPO. This discrepancy sparked sharp criticism of CONAPO. The agency, which is widely respected for its professionalism and technical expertise, stands by its estimates, however. It maintains that the surplus of birth certificates issued does not reflect "real people".

Figure 2: Coverage Errors – Exceeding Universality



Sources: Conapo estimates (based on the updated demographic model after the Conteo de Población y Vivienda 2005) and Estadísticas Vitales

Evidence for the reasons behind double registration is anecdotal, but it speaks to fundamental issues with the quality of the data, and thus the state's ability to determine whether somebody should in fact be issued a birth certificate. In Mexico, a validated copy of the birth certificate is often required for official procedures and formalities. This means that adults who interact regularly with the bureaucracy need to periodically get their certificate re-issued. One respondent mentioned that, after his family migrated from one of the Southern to one of the Northern states in Mexico, it was easier for him to just get a new certificate, listing the current state of residence as the state of birth, than to go through the trouble of periodically returning to his state of birth to obtain re-issued documents. Apparently, certain registrars were known for "being flexible" with the rules, and willing to issue new certificates.

Even though some bureaucrats might be willing to bend the rules (often in return for additional payments), trying to obtain documents to which citizens are legally entitled can still be a challenge. This is particularly true for registrations outside the legal timeframe, which often incur penalties, and require the provision of additional documentation. In the Federal District, about 20% of birth certificates have recently been issued to adults registering for the first time (Figuroa Campus 2011). Despite the apparent increase in access, one interview partner told me that for undocumented adults, it is still often easier to buy a birth certificate with another name and birthday on the black market than to go through the onerous process required for obtaining a certificate with their true identity for the first time.

An alternative explanation for the surplus that has raised concerns among public officials is the possibility that migrants arriving from Central America register their children in the Southern border states, especially Chiapas, to obtain Mexican citizenship for them. While it seems unlikely that cross-border migration would account for all or even a large part of the surplus, the fact that officials in Mexico City worry about the issue is illustrative of a fundamental problem: registering authorities do not or cannot verify the individual and familial identities of those registering. Double or even false registration is so pervasive that it calls into question the validity of the document that grants citizenship.

The previous paragraphs show that the quality of birth records is eroded by under- as well as over-registration, suggesting substantial errors of coverage. The true extent of under- and over-registration is impossible to know, as these aggregate comparisons essentially assume that every registration until universality is valid. Double or over-registration makes it impossible to estimate the pervasiveness of under-registration and, vice versa, under-registration makes it impossible to adequately assess the extent of double registration. The figures also show that the gap between the census estimate and births registered grows over time, indicating that delayed registration contributes significantly to the problem.

Analyzing Birth Records: Micro-level Analysis

To further probe the ability of the Mexican state to identify citizens, this paper draws on an original dataset of roughly 80 million individual records of births registered between 1985 and 2014. Anonymized datasets for all births registered in a given year are compiled on the basis of state-level data by Mexico's Ministry of Public Health. While names for children and parents have been removed, the records still contain basic demographic information about the date of birth, date of registration, gender as well as geographical identifiers for states, municipalities, delegations or (in the case of Oaxaca) districts. The raw datasets are organized by year of

registration (rather than year of birth), so to compile the full dataset it was necessary to harmonize the coding of all variables across time and space.[5]

At the outset of this analysis, two caveats are in order. First, the dataset contains only records of births that were registered, which is a substantial limitation especially for the early years of the analysis, during which under-registration is pervasive. Nevertheless, comparing available information about those registered to broader demographic patterns allows us to develop at least a basic notion of those who "are born and die without leaving a trace in any legal record or official statistic" (Setel et al 2007: 1). Second, there is generally no way to assess whether any specific record constitutes a duplicate or is incorrect.[6] Errors of coverage or content only become visible in the aggregate, when patterns emerge that are implausible from a demographic perspective.

To better understand who remains "undocumented", it is helpful to compare the gender distribution in the data to broader demographic patterns. According to census data, there are slightly more women in the overall population than men. Based on data from the 2000 census, the proportion of women for all age groups among the Mexican population is .512.[7] The gender distribution for all records in the dataset, which contains birth years ranging from 1900 to 2014, is .497. This gap provides a first indication that women are more likely to be undocumented than men. Since the higher proportion of women in the overall population is partially explained by the out-migration of men, for instance to work in the US, it does not provide conclusive evidence for a gendered nature of birth registration, though.

To explore the inclusion of women further, Figure 3 calculates the proportion of women in the dataset by year of birth and year of registration. The proportion of women born (and registered) between 1985 and 2014 is around .492, and fairly stable throughout the time period. It is well below the proportion in the overall dataset, though, which is indicated by the horizontal solid line. The proportion of females by year of registration, by contrast, is extremely uneven, and peaks sharply after 2000. From a demographic perspective, a sudden, sharp increase in the proportion of women in a population is unlikely, and the fairly smooth line by year of birth also rules out a sudden increase in baby girls born as a plausible explanation. Instead, the spike indicates that females who had previously been undocumented are now being registered. The sudden spike thus provides evidence for the gendered nature of delayed registration. Why are women starting to register, and girls to be registered, at higher rates after 2000?

The spike after 2000 coincides with the roll-out of the conditional cash transfer program Progresa/Oportunidades, which provides benefits to families, with a specific emphasis on women. As highlighted above, the program's technocratic, need-based orientation changed

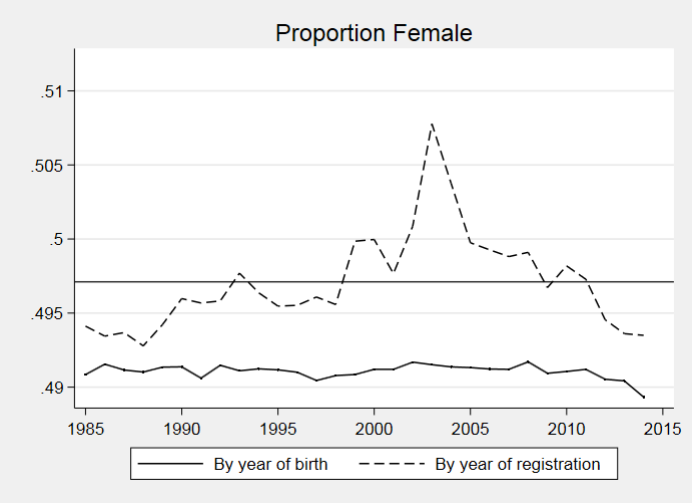
[5] Raw data from the Secretaría de Salud are not consistent in terms of variables and variable codes across years. Documentation is often minimal or incomplete. Cleaning this data, also in light of the large size of the dataset, has therefore turned out to be very time-consuming. So far, only the geographic and time variables as well as the gender variable have been cleaned and checked for consistency. Mexico changed the format and the categories on the certificate during the time period. The roll-out of the new form took several years and occurred at different speeds across states and municipalities. This adds to the challenge.

[6] In a few instances, it is obvious that a record is incorrect. There are about 25,000 cases in which the date of registration is prior to the date of birth, which then generates a negative age at registration. But even in these cases, it is most likely that the registrar simply noted the wrong year of registration when filling out the form.

[7] Calculation based on data from the UN Demographic Yearbook (https://unstats.un.org/unsd/demographic/products/dyb/DYBcensus/V1_Table1a.pdf).

information requirements for the state. Moreover, as Hunter and Brill (2016) have highlighted for Brazil and Bolivia, with the expansion of CCTs official documents, like birth records, acquired a concrete, tangible value for marginalized populations. While evidence for the link between the roll-out of social programs and an increase in women registering remains indirect, further graphs, not shown here, bolster confidence in the link. The graphs examine the changing proportion of females across states, and show that the spike after 2000 is most pronounced in relatively poor and rural states, like Guerrero and Oaxaca. These states received significant funding under Oportunidades. Moreover, the spike emerges much more sharply among registrations delayed beyond the 12 months window. Delayed registration therefore appears to particularly affect women.

Figure 3: Gender Distribution Among Registered Births



To further probe the gendered nature of delayed registration, Figure 4 examines the average age at registration in months by gender. Even though the lines for males and females track each other fairly closely over time, the delay is always longer for females. Peaks in the average age at registration suggest that more older citizens are obtaining birth records for the first time.[8] We would expect to see the demand for birth records among older cohorts to rise if the relevance and value of documents for these citizens increases. Two spikes are noteworthy. The first is the peak after 2000, which also emerged in the previous graph. The distance between the two lines widens after 2000, which again speaks to the role of social policies aimed at female beneficiaries in expanding birth registration among women. The second, smaller spike occurs in the mid-1990s. This first rise in the demand for records coincides with the establishment of a national voter registry under the auspices of the Federal Electoral Institute (IFE). The IFE was created in 1990 in the aftermath of the 1988 federal elections, which were marred by substantial irregularities and widespread allegations of fraud. IFE, which has now been transformed into the National Electoral Institute (INE), is widely credited with cleaning up the electoral process in Mexico, and paving the way for a peaceful, electoral transfer of power from the hegemonic PRI to opposition parties. The IFE created and maintained a national registry of voters, and the voter ID issued by IFE (*credencial para votar*) serves as the equivalent of a national picture ID card. While more in-depth analysis is needed, the rise in birth registration after 1990 suggests

[8] Or second time, in the case of double registration, but there is no way to distinguish real first-time registrants from those who have been registered before and misplaced their documents.

that political and electoral dynamics might also play a role in increasing the demand for birth records. [9]

Figure 4: The Gendered Nature of Delayed Registration

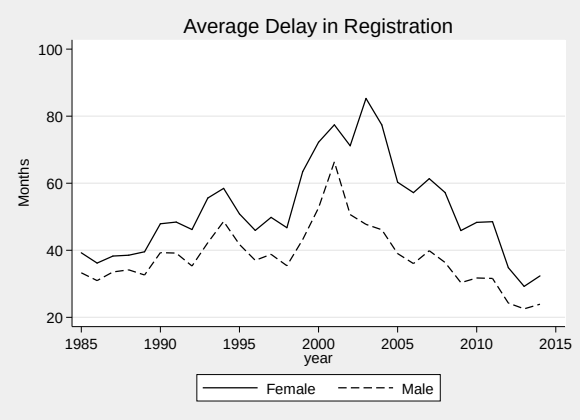
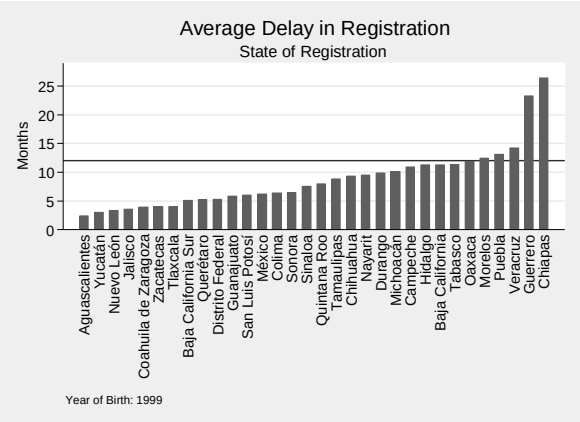
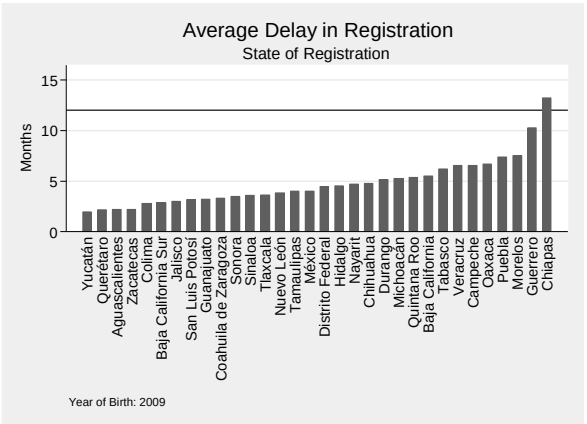


Figure 5: Territorial Unevenness in Delayed Registration



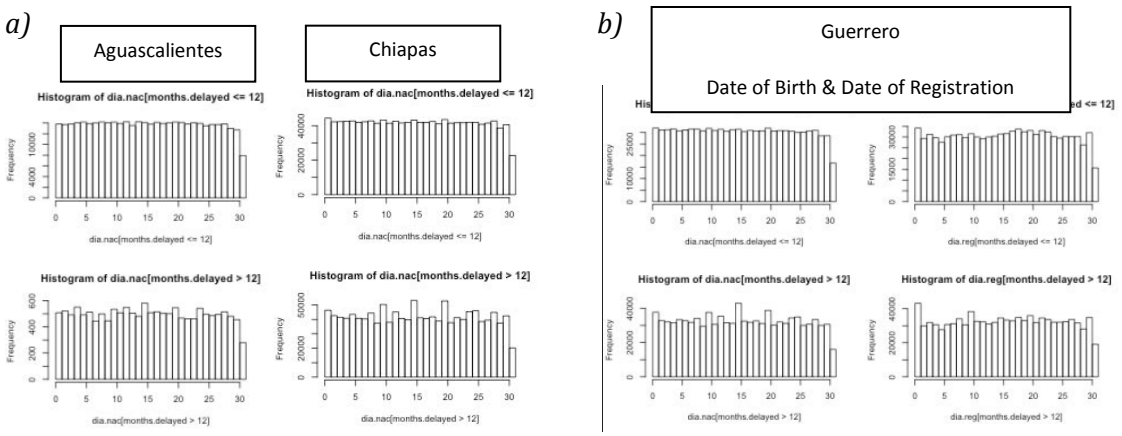


The aggregate analysis above suggests that the incidence of coverage errors differs across states, and that errors are compounded by delays in registration. To further explore this link, Figure 5 plots the average delay in registration per state for 1999 and 2009. The comparison of the two years indicates progress towards timely registration, which is consistent with INEGI's (2011) analysis of aggregate data. Nevertheless, it is important to keep in mind that the dataset of birth records only includes births registered until 2014, so for the 2009 data comparatively less time has passed in which delays could pile up. Against this background, it is noteworthy that the average age at registration in Chiapas still exceeds 12 months. Moreover, those states with the highest delays in Figure 5 overlap with those where Figure 1 suggests coverage errors.

[9] One way I might be able to get at this as I develop this project further is to analyze subnational variation in the timing and degree of the entry of older cohorts across states and municipalities. The rise in electoral competitiveness was spatially uneven. Since trying to obtain records for the first time as an adult is difficult, it appears to be a classic incidence of a situation where party brokers can help loyal constituencies navigate the bureaucracy. As state-level elections become more competitive, we may therefore see an increasing interest in birth records in relatively poor, rural municipalities which have historically been bastions of support for the formerly hegemonic PRI.

As a final note, it is worth pointing out that delays in registration not only facilitate coverage errors, they also go hand in hand with content errors. The reasons are twofold. First, the more time has passed between birth and day of registration, the less likely registering guardians are to recall the exact date of birth. For citizens registering as adults, the exact date is almost certainly guess-work, unless somebody in the family kept records. And second, for societal groups that do not have access to timely registration, exact dates likely possess less relevance to begin with, and are therefore less likely to be remembered.[10] An illustration of the connection between delays and content errors is provided by the (still rather rough) histograms in Figure 6a, which plot reported date of birth for Aguascalientes and Chiapas, the top and bottom performer with regard to the average delay in registration. The top panel captures timely registrations within the first 12 months of a child's life. The bottom panel reflects registrations beyond the first 12 months, and shows considerable evidence for heaping, even in Aguascalientes. Figure 6b, for the state of Guerrero, shows the same pattern for dates of birth on the left. In addition, however, it also shows heaping with regard to the day of registration (right), with considerably more delayed births registered on the first day of the month. This suggests more serious issues with how diligently forms are filled out by registrars.[11]

Figure 6: Heaping



Conclusions and Implications

Birth records offer a unique lens for examining the ability of the Mexican state to identify citizens. The analysis shows that birth records are simultaneously affected by the under-registration of some citizens and the duplication of others. These errors are driven, in part, by the often substantial delays between the date of birth and date of registration. Delays make it extremely difficult for officials to verify individual and familial identities, and thus to ensure that the reported information is correct. Errors of content, such as a type of misreporting called heaping, are also compounded by delays of registration.

[10] The notion that dates of birth do not necessarily have to be accurate, though, is not necessarily limited to poor and rural communities. One friend, from a middle class family, mentioned that her grandfather registered all his children as having been born on national holidays, such as the Day of the Mexican Revolution or Independence Day to express his patriotism.

[11] Graphs not shown here also provide considerable evidence for heaping by birth year, and how that heaping increases with delays.

Methodologically, the paper draws heavily on tools from demography. A fair questions would therefore be why the quality of birth records is of concern to political scientists. Why not leave the analysis of these records in the capable hands of demographers? The reason is that birth records, even more than other types of information collected by states about populations and terrain, speak to the relationship between the state and citizens. Errors of coverage and content are not only statistical issues, but have fundamental bearing on the nature of citizenship. Birth records confer rights to their holders, and are the prerequisite for claims based on nationality or age. Where citizens do not seek these documents, or are unable to obtain them, this effectively renders all of these rights moot. States then either provide so little for some groups of citizens, or fail to uphold their rights so systematically, that official documents have little value. An examination of birth records therefore sheds light on unevenness in the intensity of citizenship (O'Donnell 1993) across segments of the population.

From this perspective, the analysis of birth records in Mexico provides both good and bad news. On the positive side, it appears that coverage has increased, particularly among poor, rural populations and women, two groups that are among the most likely to be undocumented in their own country. This expansion of access to documents has been aided substantially by the roll-out of social programs aimed at these groups. In this sense, it suggests a transactional conceptualization of citizenship, where a state project to implement better targeted poverty alleviation, and the prospect of material benefits for marginalized citizens, drives an expansion of access to birth registration in terms of supply and demand (Hunter and Brill 2016).

This expansion of access is no small feat. Providing these groups with documents, and therefore intensifying their citizenship and relationship with the state through social policies not only improves their material position in the short term. It also raises the hope that children growing up in families covered by these programs, who now possess official documents to prove their identity, will later be able to access other benefits that are supposed to come from the protection of identity. For these families and especially for their children, having documents therefore at least potentially opens up doors that were previously closed.

Nevertheless, the analysis in this paper is also very much a cautionary tale about the quality of records. For the case of Mexico, the analysis calls into question the usefulness of birth certificates as a way for the state to pinpoint individual citizens. Long delays between date of birth and registration weaken the state's ability to verify identities. The expansion of access to documents has gone hand in hand with an apparent increase in registrations beyond universality, even though under-registration most likely still continues. While it has been suggested that some of the surplus registrations might be fraudulent, honest mistakes by citizens, and their best efforts to navigate an unresponsive and ineffective bureaucracy, most likely account for much of the surplus. The state, however, cannot effectively use these records to identify individual citizens. The quality of records varies geographically and across societal groups. This has two important implications.

First, "scaling down" (Snyder 2001) to the subnational level highlights how subnational variation in governance and state capacity contribute to a territorial unevenness in citizenship. While federal laws determine criteria for the conferral of citizenship, state civil codes and the responsiveness of state bureaucracies influence the relative ease with which those entitled to citizenship can claim their rights. This has implications beyond Mexico. In the US, for instance, rules for replacing birth certificates and other identity documents vary widely across states.

There are charities specifically devoted to helping citizens replace documents that they have lost through evictions, homelessness or just bad luck.[12]

Second, "scaling back up" to the national level reveals that the poor quality of records in some parts of the country is not a geographically isolated issue. Citizens move around within their country, and if citizens born in one state have a hard time obtaining records to which they are entitled, this affects their status even after they move. Similarly, if records from one part of the country are poorly verified, this has implications for the ability of other states to monitor access to public services. Internal migration, especially towards urban areas, has been considerable in Mexico. Since citizenship is national a failure of the state "in even the most remote parts of the country can affect the state in the capital city by denying state components there resources and support from the larger society" (Migdal, Kohli, and Shue 1994).

[12] Bradley, Patrick Marion. "The Invisibles: The cruel Catch-22 of being poor with no ID", Washington Post Magazine, published on June 15th, 2017. (Available at: https://www.washingtonpost.com/lifestyle/magazine/what-happens-to-people-who-cant-prove-who-they-are/2017/06/14/fc0aaca2-4215-11e7-adba-394ee67a7582_story.html?utm_term=.0a32b02581d7).

References (not yet complete)

D'Arcy, M. and Nistotskaya, M., 2017. State first, then democracy: Using cadastral records to explain governmental performance in public goods provision. *Governance*, *30*(2), pp.193-209.

Diaz-Cayeros, A., Estévez, F. and Magaloni, B., 2012. Strategies of vote buying: democracy, clientelism and poverty relief in Mexico. *Unpublished manuscript, Stanford University*.

Figueroa Campos, B., 2011. El registro de los nacimiento en el Distrito Federal, características y problemas de la información derivada de estos registros. In: Valdés, L.M., Conmemoración del 150 aniversario del Registro Civil. *Fundamentos y Reflexiones, Universidad Nacional Autónoma de México, Instituto de Investigaciones Jurídicas, Estudios Jurídicos*.

Grech, V., Savona-Ventura, C. and Vassallo-Agius, P., 2002. Unexplained differences in sex ratios at birth in Europe and North America. *BMJ*, 324(7344), pp.1010-1011.

Hunter, W. and Brill, R., 2016. Documents, Please. *World Politics*, *68*(2), pp.191-228.

INEGI, UNICEF. *Derecho a la identidad: la cobertura de registro de nacimientos en México en 1999 y 2009*.

Lee, M.M. and Zhang, N., 2017. Legibility and the Informational Foundations of State Capacity. *The Journal of Politics*, *79*(1), pp.118-132.

Mahapatra, P., Shibuya, K., Lopez, A.D., Coullare, F., Notzon, F.C., Rao, C. and Szreter, S., 2007. Civil registration systems and vital statistics: successes and missed opportunities. *The Lancet*, *370*(9599), pp.1653-1663.

Migdal, J.S., Kohli, A. and Shue, V. eds., 1994. *State power and social forces: domination and transformation in the Third World*. Cambridge University Press.

O'Donnell, G., 1993. On the state, democratization and some conceptual problems: A Latin American view with glances at some postcommunist countries. *World Development*, *21*(8), pp.1355-1369.

Scott, J.C., 1998. *Seeing like a state: How certain schemes to improve the human condition have failed*. Yale University Press.

Setel, P.W., Macfarlane, S.B., Szreter, S., Mikkelsen, L., Jha, P., Stout, S., AbouZahr, C. and Monitoring of Vital Events (MoVE) writing group, 2007. A scandal of invisibility: making everyone count by counting everyone. *The Lancet*, *370*(9598), pp.1569-1577.

Snyder, R., 2001. Scaling down: The subnational comparative method. *Studies in Comparative International Development (SCID)*, *36*(1), pp.93-110.

Soifer, H.D., 2013. State power and the economic origins of democracy. *Studies in Comparative International Development*, *48*(1), pp.1-22.

Valdés, L.M.(ed.), 2011. *Conmemoración del 150 aniversario del Registro Civil*. *Fundamentos y Reflexiones,* Universidad Nacional Autónoma de México, Instituto de Investigaciones Jurídicas, Estudios Jurídicos.