WPS6884

Policy Research Working Paper 6884

Rewarding Provider Performance to Enable a Healthy Start to Life

Evidence from Argentina's Plan Nacer

Paul Gertler
Paula Giovagnoli
Sebastian Martinez

The World Bank
Latin America and the Caribbean Region
Health, Nutrition and Population Unit
May 2014

Policy Research Working Paper 6884

Abstract

Argentina's Plan Nacer provides insurance for maternal and child health care to uninsured families. The program allocates funding to provinces based on enrollment of beneficiaries and adds performance incentives based on indicators of the use and quality of maternal and child health care services and health outcomes. The provinces use these resources to pay health facilities to provide maternal and child health care services to beneficiaries. This paper analyzes the impact of *Plan Nacer* on birth outcomes. The analysis uses data from the universe of birth records in seven Argentine provinces for 2004 to 2008 and exploits the geographic phasing in of *Plan Nacer* over time. The paper finds that the program increases the use and quality of prenatal care as measured by the number of visits and the probability of receiving a tetanus vaccine. Beneficiaries' probability of low birth-weight is estimated to be reduced by 19 percent. Beneficiaries have a 74 percent lower chance of in-hospital neonatal mortality in larger facilities and approximately half this reduction comes from preventing low birth weight and half from better postnatal care. The analysis finds that the cost of saving a disability-adjusted life year through the program was $814, which is highly cost-effective compared with Argentina's $6,075 gross domestic product per capita over this period. Although there are small negative spillover effects on prenatal care utilization of non-beneficiary populations in clinics covered by Plan Nacer, no spillover is found on their birth outcomes.

This paper is a product of the Health, Nutrition and Population Unit, Latin America and the Caribbean Region. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org.
The authors may be contacted at gertler@haas.berkeley.edu and pgiovagnoli@worldbank.org.

Rewarding Provider Performance to Enable a Healthy Start to Life: Evidence from Argentina's Plan Nacer

Paul Gertler, Paula Giovagnoli, and Sebastian Martinez

JEL Classification: I12, I13, I15, I18

Keywords: Pay-for-performance, results-based financing, provider performance, birth outcomes, impact evaluation, maternal and child health

Author Affiliation: Paul Gertler (Gertler@haas.berkeley.edu) is the Li Ka Shing Professor at the University of California, Berkeley. Paula Giovagnoli (pgiovagnoli@worldbank.org) is an economist with the World Bank. Sebastian Martinez (smartinez@iadb.org) is a lead economist at the Inter-American Development Bank.

Acknowledgments: This paper has benefited greatly from assistance from the Área Tecnica Plan Nacer, the Área de Capitas Plan Nacer, and the REMEDIAR program of the Ministry of Health, Argentina. The report has also greatly benefited from comments provided by the Ministry of Health officials including Maximo Diosque, Martin Sabignoso, Humberto Silva, Álvaro Ocariz, Ricardo Izquierdo, Ana Sala, Ramiro Flores Cruz, Santiago Garriga, and Leonardo Zanazzi. The paper has also benefited from comments by World Bank staff including Daniela Romero, Vanina Camporeale, Gaston Blanco, Margaret Grosh, Mario Navarrete, Luis Perez, Nazmul Chaudhury, and Christel Vermeersch. The authors acknowledge excellent research assistance from Rafael Ramirez, Pablo Celhy, Adam Ross, Sigrid Vivo, and Joshua Gruber. Finally, the authors gratefully acknowledge financial support from the World Bank Strategic Impact Evaluation Fund (SIEF), and The World Bank Health Results Innovation Trust Fund (HRITF). The content and conclusions expressed in this paper are those of the authors themselves and do not necessarily represent the views of the World Bank, Argentine National Ministry of Health, or the funders of the research. All errors and omissions are the responsibility of the authors alone.

1. INTRODUCTION

Access to quality maternal and child health care services for the poor is a major policy concern in almost all countries globally. A healthy birth is a critical initial condition for long-term success in life. Low birth weight in particular is associated with poor cognitive development, persistent lifelong health problems, low school achievement, and reduced lifetime earnings (Behrman and Rosenzwieg 2004; Almond et al. 2005; and Black et al. 2007). Access to quality maternal and child health services is especially important for the poor as children born into poverty are much more likely to have low birth weight (Currie 2011). In this paper we examine the impact on birth outcomes of Argentina's Plan Nacer, an innovative program that provides health insurance for maternal and child medical care to poor households and uses performance payments based in part on health outcomes to incentivize the provision of high-quality care to beneficiaries.

In many countries the responsibility for ensuring access to health services for the poor has been devolved to local governments but is cofinanced with the national government. For example, Medicaid in the United States, Seguro Popular in Mexico, and Plan Nacer in Argentina are all managed at the province level but receive supplemental funding from their federal counterparts. In the cases of Medicaid and Seguro Popular, the amount transferred to states is a function of the number of beneficiaries enrolled.

In contrast, Plan Nacer uses an innovative pay-for-performance model (P4P) that provides incentives to the provinces to improve health outcomes by conditioning the financing not only on enrollment but also on the achievement of a specific set of indicators that include health outcomes. The provincial Plan Nacer programs then pass these incentives on to health clinics and hospitals by paying them for beneficiary use of maternal and child medical services at a quality indicated by the provision of clinical services that are appropriate to the purpose of a patient's visit.

The novel incentive structure in Plan Nacer is designed to address long-standing quality problems in the public provision of health services by making providers more accountable (Chaudhury et al. 2006; Das et al. 2008; and World Bank, 2004). Such linking of finance to local government performance has been used in Indonesia, where financial incentives are paid to local governments for improvements in village-level health and education outcomes (Olken et al. 2013) and in India where local governments are paid if they eliminate open defecation at the village level (Spears 2013). Many more governments use performance incentives at medical care–provider levels. P4P schemes for health care providers were introduced in high-income countries beginning in the 1990s and in low- and

middle-income countries in the last 10 years; however, few of them condition payments on health outcomes (Miller and Singer 2013).[1]

In this paper we evaluate the impact of Plan Nacer on birth outcomes and neonatal mortality. Our identification strategy exploits the geographic phasing of the program across clinics over time. We use difference-in-difference models that allow us to control for time-invariant confounding factors that may exist between population groups receiving care from clinics incorporated into the program early compared to those included later, and time-varying factors within provinces that are common to both early and late adopters. Specification tests, used to evaluate pre-intervention trends, support a causal interpretation of our results from difference-in-difference models.

Our analysis uses information obtained from birth certificate records in pubic maternity hospitals combined with administrative data sources for the seven provinces in northern Argentina for 2004 through 2008, covering the first phase of Plan Nacer scale-up. Birth certificates provide accurate data on prenatal care, birth outcomes, and in-hospital neonatal mortality. The size of the dataset allows us to investigate the impact of Plan Nacer on relatively rare events such as low birth weight and neonatal mortality, which is typically not possible in surveys with smaller samples.

Overall, these results show that Plan Nacer was effective in improving birth outcomes and reducing neonatal mortality, and that it achieved these gains through expanded access and improvements in the quality of health services in the public sector. Specifically, the program is associated with a significant increase in the number of prenatal care visits and the quality of prenatal care and delivery, measured by the probability of receiving a tetanus toxoid vaccine and a reduced probability of birth by cesarean section. We also find that beneficiary newborns had substantially better health outcomes in terms of less low birth weight and reduced neonatal mortality. Specifically, we find a 19 percent reduction in the probability of low birth weight, and a 74 percent reduction in neonatal mortality.

In addition to direct program impact, we also analyze program spillovers onto nonbeneficiary populations — that is, the impact of Plan Nacer on a nonbeneficiary who obtained care at a clinic incorporated into the program. Positive spillovers could be expected if program resources are invested in equipment, supplies, and system improvements in quality of care that affect both beneficiary and nonbeneficiary patients. On the other hand, we might expect negative spillovers

1. Such linking of finance to performance has also been used in state and local school reform in the United States (Race to the Top, No Child Left Behind) as well as incentive payments for teachers (see, for example, Muralidharan and Sundararaman 2011).

if providers reallocate effort and resources to beneficiaries away from nonbeneficiaries, that is, a "crowding out" of non-beneficiaries. Our analysis suggests some evidence of minor crowding out in the use and quality of health services, in particular small reductions in the number of prenatal care visits and increases in the probability of cesarean section at birth. However, these some reductions in care were not large enough to affect non-beneficiary birth outcomes.

This paper contributes to a relatively small literature on P4P in health care in low- and middle-income countries. Despite the expansion of P4P in health schemes around the world, empirical evidence as to their effectiveness remains limited, particularly in low- and middle-income countries (Miller and Singer 2013). There are to our knowledge, four well-identified related evaluations in other low- and middle-income countries. In Rwanda, P4P payments to primary health care clinics resulted in improved access to higher-quality care and improved child health outcomes (Basinga et al. 2011; Gertler and Vermeersch 2013). In Indonesia, performance incentives to villages for improvements in health and education outcomes led to an improvement in nutritional outcomes (Olken et al. 2013). Miller et al. (2012) found that bonus payments to schools significantly reduced anemia among students in China. Finally, incentives to village leaders in India to eliminate open defecation led to reductions in stunting and child mortality (Spears 2013).

2. PLAN NACER

2.1. Origins

Historically, Argentina has provided services to the uninsured through its universal public health care system. The public system offers free health services to all Argentines and foreigners, independent of whether the individual has formal insurance or not. Formal insurance is typically provided through employers (Obra Social) or purchased from private insurance companies. Those covered by formal insurance receive care primarily through private providers and the social security or Obra Social system. The uninsured, on the other hand, include the unemployed, those working outside the formal sector, and those unable to afford private insurance. Thus, in practice, the poor rely heavily on the publicly financed system.

The public health care system continues to care for a large share of the population; approximately 14.95 million people, or 38 percent of Argentina's population, were uninsured in 2009. Of these, 2.07 million or approximately 14 percent are eligible for Plan Nacer based on age or pregnancy status. The responsibility for the provision of health services to this population rests mainly with the country's 23 provinces and the Autonomous City of Buenos Aires. Public health services are financed through fixed historical budgets, which are often

insufficient to meet the local demands for services, and provide little to no flexibility for local service providers on how resources are used (Musgrove 2010).

Plan Nacer was launched in 2004 following the deterioration in maternal and child health indicators resulting from the 2001 economic crisis. Plan Nacer is designed to improve the health status of uninsured pregnant women and children by channeling more resources to the public health care system and creating incentives to use those resources more efficiently. The program covers women during pregnancy and up to 45 days after birth (or the loss of the fetus). Coverage for children is up to age six and concentrates on services during the first year. All other care outside of the Plan Nacer package of benefits must be covered by regular provincial health services.

The program was first launched in nine provinces in northern Argentina and then expanded to cover the rest of the country beginning in 2007. By 2012 the program was operating nationally with high coverage levels.[2] The nine northern provinces where Plan Nacer was initiated have the highest levels of maternal and child mortality and experienced the most severe decline in health indicators during the 2001 crisis. The provinces have a population of 8.26 million[3] and the uninsured population of women and children in 2008 was estimated at 68,514 women and 569,984 children.[4] By December 2008, the program's enrollment had reached 82 percent of the eligible population, a total of 1,886 contracted health providers (57.4 percent of all providers), and transferred over $81 million[5] (Musgrove 2010).

2.2. Payment to Provinces

Plan Nacer was designed to supplement the existing public financing system with an innovative P4P model that incentivizes the provision of quality priority maternal and infant health services. The program is overlaid on the existing national and provincial health financing structures, and represents additional funding beyond the historical administrative budgets. Through Plan Nacer, the national government reimburses provinces on a per capita basis at a maximum cost of $8 per person per month. The provinces receive $5 (60 percent of the maximum per capita payment) for every eligible individual enrolled in the program and up to an additional $3 (40 percent of the maximum payment) if health targets for the eligible

2. In 2012 the program was expanded to include additional beneficiary populations and prioritized services under the name "SUMAR."
3. Argentina. National Institute of Statistics and Census 2010.
4. Argentina. National Ministry of Health. Secretary of Health Promotion and Programs. Plan Nacer Central Implementation Unit (Unidad Ejecutora Central [UEC]), 2008 a.
5. Argentina. National Ministry of Health. Secretary of Health Promotion and Programs. Plan Nacer Central Implementation Unit (Unidad Ejecutora Central [UEC]). "Informes de Gestión: Plan Nacer — 2004– 2009." Buenos Aires. Argentina.

population are achieved. Thus, the program provides explicit incentives to enroll the target population of uninsured mothers and children, and to provide services that improve health outcomes of the eligible population.

Health targets are measured using ten specific indicators called tracers (table 1). The tracers are derived from best practice clinical protocols. The tracers include health outcomes such as share of newborns with low birth weight (that is, less than 2,500 grams) and APGAR scores greater than 6. They also include the use of priority services such as beginning prenatal care in the first 20 weeks of pregnancy, receiving Venereal Disease Research Laboratory (VDRL) test and tetanus vaccines prior to delivery, reproductive health counseling within 45 days of delivery, on-time and complete child immunization, and complete and on-time well-baby visits.

Table 1: Tracers

	Tracer	Description
1.	Early detection of pregnancy	First prenatal checkup before week 20 of pregnancy
2.	Effectiveness of obstetric care	Five-minute APGAR scores of over 6
3.	Effectiveness of prenatal care	Birth weight of over 2,500 grams
4.	Effectiveness of prenatal care	Mother given VDRL test and tetanus toxoid vaccination prior to delivery
5.	Proper cause of death review of infant and maternal deaths	Proper cause of death review of infant and maternal deaths
6.	Immunization coverage	Children under 18 months of age given measles vaccine or triple vaccine
7.	Sexual and reproductive health	Mothers receive counseling within 45 days of delivery
8.	Well-child follow-up until one year of age	Children <1 with complete record of preventive checkups
9.	Well-child follow-up from one to six years of age	Children 1 to 6 years old with complete record of preventive checkups
10.	Inclusion of the indigenous population	Providers with staff trained to provide care to indigenous population

Source: Manual Operativo (2005)

Performance targets associated with the tracers for each province are set with the provinces in the annual agreement between the parties and are measured

using national statistical sources. The performance payment is divided equally among the ten tracers, with four percentage points assigned to each, totaling up to 40 percent. If the target is met, the province receives the full 4 percent of the capita for that tracer. If it does not meet the target, it receives nothing for that tracer.[6] Payments are made every four months.

2.3. Payment to Clinics

Each province uses these resources to pay clinics for priority maternal and infant health services on a fee-for-service basis. The 80 services included are defined at the national level, with the provincial health authority determining the prices of the services. Only clinics affiliated with the program can be paid and only for services included in the package provided to persons enrolled in the program. The program includes an intensive process for auditing and verifying clinic records to ascertain the validity of payments made.

The services in the nomenclature package were chosen based on their priority in terms of maximizing the probability of good maternal and child health outcomes. Services include preventive and diagnostic visits for pregnant women, new mothers and children, delivery care, laboratory analyses and imaging, and health care promotion (table 2). Some services include visits such as prenatal and well-baby care; other clinical services provided during these visits, such as vaccines, diagnostic tests, and reproductive health counseling included in the package, are meant to represent higher quality of care. Hence, clinics are paid more for providing higher-quality care to more beneficiaries.

Plan Nacer resources supplement the regular operating budget of the clinic. General guidelines for the use of resources by providers are set at a national level, and provinces are allowed to impose additional restrictions for service providers in their jurisdictions. However, within these guidelines, resources may be used at the discretion of the provider to improve the quality of health services. Specifically, expenditures permitted by the program include medical and office supplies, maintenance (infrastructure), investments (infrastructure and medical equipment), agreements for works and services (staff recruitment), and staff incentives.[7]

6. In May 2008, the "all or nothing" scheme was modified to a sliding scale with four thresholds rather than one (Argentina. Ministry of Health. 2008b).
7. For more information on the structure and operation of the program, see Argentina, Ministry of Health 2008b.

Table 2 Categories of Items Covered in Nomenclature

Groups	Subgroups	Practices (Some examples)
Women	*Pregnant women*	Health Education Consultation during pregnancy
		Papanicolau and Colposcopy
		Antitetanic Vaccine
	High risk pregnancy	High-risk Pregnancy consultation
		Human immunodeficiency virus (HIV) care during pregnancy
	Delivery	Delivery
		Cesarean
	Purperium	Measles immunization
		Puerperium counseling
Children	*New-born baby*	Immunization of the new-born children
		Incubator up for a period of 48 hours
		Immediate treatment in case of HIV vertical transmission
		Ophthalmologic consultation
	Infant under 6 years old	Follow-up consultation
		Dental care counseling
Laboratory		Pregnancy test
		Colposcopy in pregnant control
		Blood extraction
		Blood test
Images		Thoracic XR
		Ecography
Community		Detection of pregnant women in their first quarter of pregnancy by sanitary or health care agents
		Round of sanitary agent in rural area
		Socio – Epidemiologic diagnosis of population at risk
		Reunions for feeding guidelines promotion
		Infant development promotion meetings
Transport		Newborn emergency transportation service

Source: Nomenclature's Practices. Provincial Maternal -Child Health InvestmentProject. Presentation at Sixth plenary meeting of the Leading Group on 23 Solidarity Levies to Fund Development

2.4. Rollout and Scale-Up

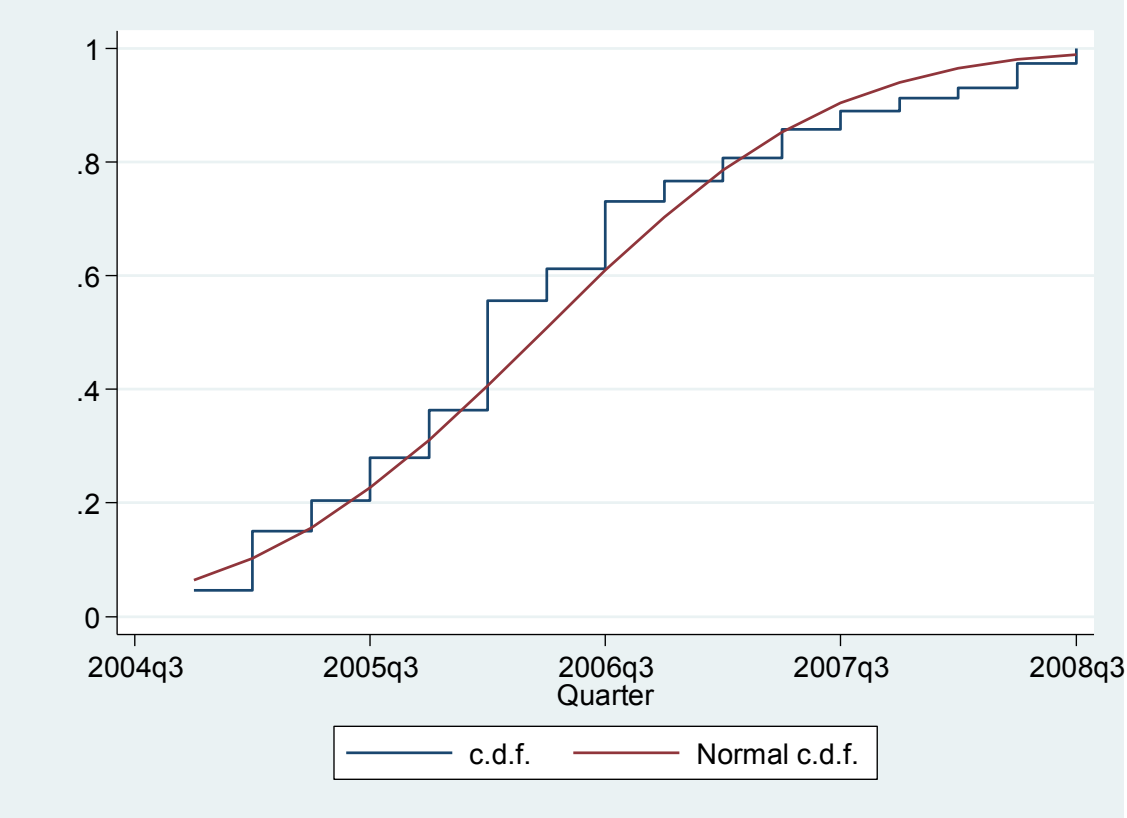
The initial phase of the program implementation focused on negotiating contracts with provinces and on the establishment of provincial implementation units. Once the program had been established in the province, the process of enrolling clinics began by entering into agreements with each participating provider. The expansion of the program's coverage in each province is determined by an annual action plan for implementation that set forth enrollment targets in terms of the number of providers and the number of eligible population that should be incorporated into the program in a given year.

Implementing the program in each province involved implementing the required systems (for example, payment for services and monitoring of results), enrolling providers, and finally, identifying and enrolling the eligible population; these multiple actions in each province resulted in a gradual rollout of the program over time. For the purposes of our analysis, we consider the program to have

started operations in an area of intervention once the clinic submitted the first bill for payment to Plan Nacer for services provided to beneficiaries.

Figure 1 plots the timing of incorporation of clinics into the program. Overall, the incorporation is smoothed over the five-year period. The first billing of services took place during the last quarter of 2004 in the province of Tucumán. The remaining provinces started billing from 2005 onwards. By the end of 2008, the proportion of clinics that had billed at least once to Plan Nacer was close to 100 percent.

Figure 1: Timing of Clinics Enrolling in Plan Nacer



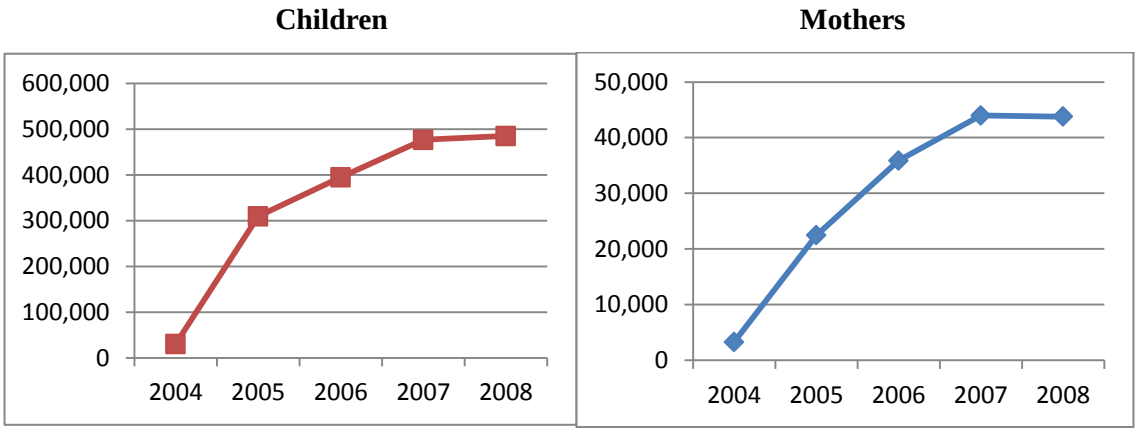
Source: Plan Nacer clinic billing and payment administrative records

Note: The figure plots the cumulative proportion of clinics incorporated by quarter-year. The quarter-year of incorporation refers to the date the clinic submitted the first bill for services provided to Plan Nacer beneficiaries.

In contrast to service provider enrollment, beneficiary enrollment of eligible mothers and children took significantly more time (figure 2). In practice, the

identification and enrollment of eligible populations in towns and cities was strongly linked to the participation of the area's provider, as providers were typically assigned this task. Once a provider was enrolled in the Plan, it was able to combine its efforts with the province's to enroll eligible women and children into Plan Nacer. Since enrollment is not instantaneous, the longer the tenure of a provider on Plan Nacer, the more potential beneficiaries the provider was able to find and enroll.

Figure 2: Beneficiary Enrollment in Plan Nacer



Source: Area de Capita. National Ministry of Health (2005-2008).

2.5. Program Expenditures

Between 2005–08 Plan Nacer was implemented in nine northern provinces, and total expenditures totaled $107 million. Plan Nacer resources supplement provincial public services financed through provincial budgets and through Obras Sociales, health insurance for those employed in the formal sector (table 3). Plan Nacer expenditures increased total public spending on health services in these provinces by 1.4 to 3.5 percent during this period. However, the percentage increase is substantially higher for maternal and child health care. Unfortunately, expenditure information broken down for maternal and child health services is not available in Argentina.

Capitation payments from national to provincial units come in two installments: (1) 60 percent of the maximum payment is disbursed monthly based on the number of verified registered beneficiaries; and (2) up to 40 percent of the maximum is transferred every four months after verification and certification that the province actually met the tracer targets. During the first eight months of the program, there existed a grace period for which payments were made 100 percent on the basis of enrollment. Figure 3 shows the total capitation payments from the

national government to the provincial governments over time (blue line) and the performance capitation associated with the traces (red line).

Table 3: Public Spending on Health Services in 9 Provinces (millions, USD) 2005–08

Years	2005	2006	2007	2008
Health services (publicly financed)	486	568	852	1134
Health services (Obras Sociales)	278	307	392	498
Total health services spending	764	875	1244	1632
Plan Nacer expenditures	**11**	**31**	**36**	**33**
Plan Nacer as % of total expenditures	**1.4**	**3.5**	**2.9**	**2.0**

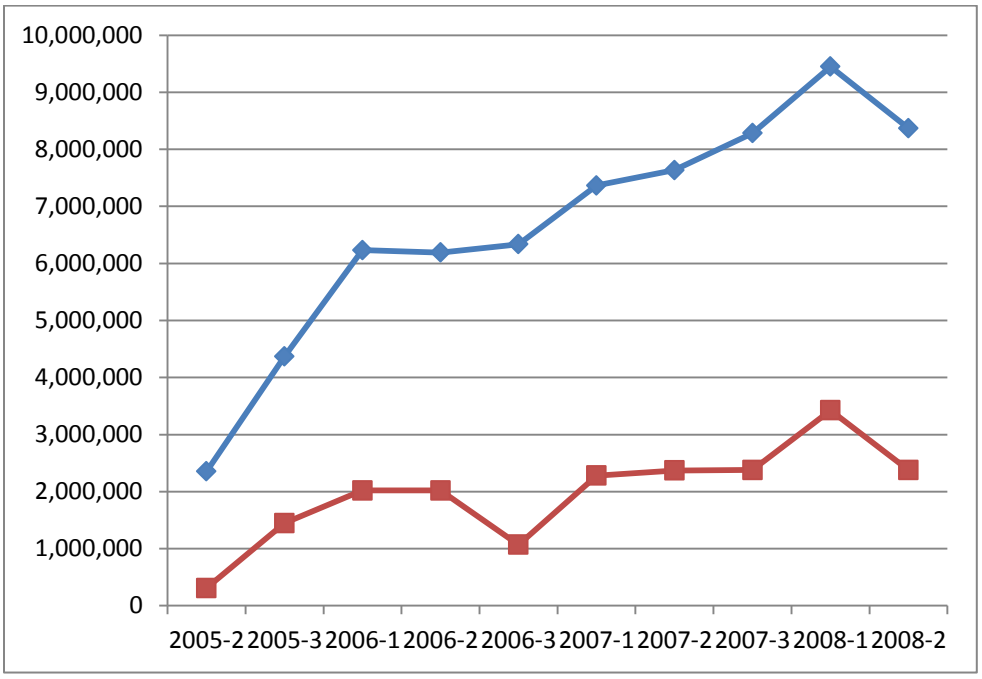
Source: Unidad de Financiamiento Internacional de Salud (UFIS), Ministerio de Salud Nacional. 2005-2008.

The provinces by and large met most but not all of the tracer targets. Figure 4 reports the share of tracer targets meet by the provinces every four months. Except for a few periods, the provinces meet between 70 and 80 percent of the targets and as a result were rewarded with most of the maximum payment possible.

Plan Nacer finances a number of activities (table 4). The largest category encompasses payments to the facilities for care services for Plan Nacer beneficiaries. Payments from provincial governments to the facilities for services accounted for $67.11 billion or 63 percent of total expenditure over the whole period. The payment-for-services share rose over time, starting at 23 percent in 2005 and rising to 71 percent in 2008, reflecting the growth in enrollment so that fixed costs are spread over a larger base. Payments to facilities represent 94 percent of capitation payments from the national government to the provinces.

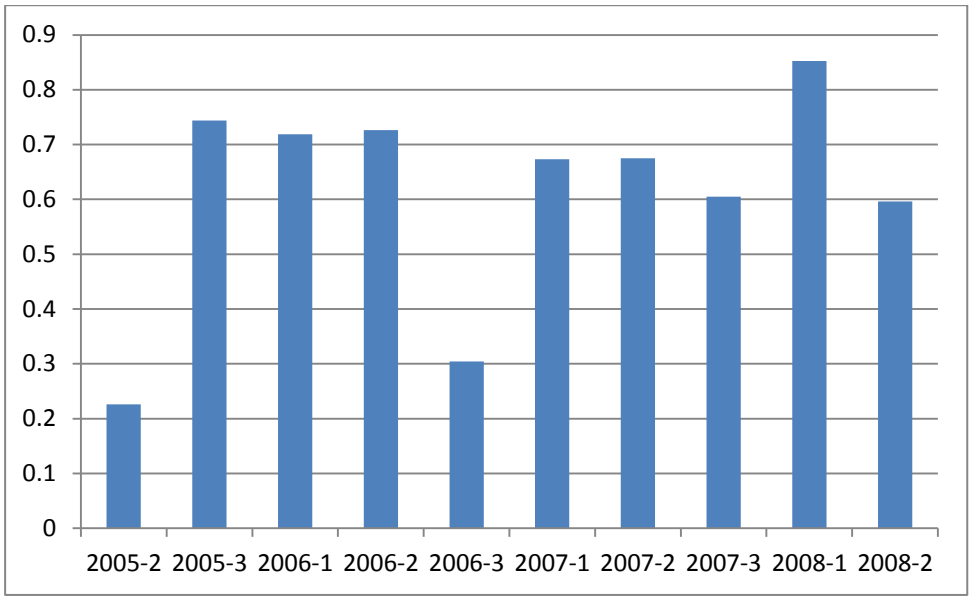
The second largest category of Plan Nacer expenditure was investment in durables such as medical equipment in the medical care facilities and in software, communications, office equipment, and vehicles in the provincial Plan Nacer offices. Investment in medical facilities was designed to ensure that the facilities were able to deliver quality care, and in the provincial program office to facilitate program implementation. Between the years 2005 to 2008, program investment accounted for about 21 percent of total expenditure. As expected, most of this investment occurred in the earlier years, representing about 30 percent of total expenditure in 2005 and falling to 8 percent in 2008.

Figure 3: Total Payments from National Government to Provinces (Blue) and Performance Payments from Meeting Tracer Targets (Red)



Source: Informes de Gestión: Plan Nacer (2005-2008).

Figure 4: Share of Tracer Performance Targets Achieved



Source: : Informes de Gestión: Plan Nacer (2005-2008).

The remaining expenditure covers overhead costs including (i) external auditing and verification of indicator reporting and payments at both the provincial and facility levels and for upgrading and maintaining the information system; (ii) training; (iii) national administration; and (iv) some one-time front-end fees. While over the whole period, overhead costs accounted for 16 percent of expenditures, by 2008 they accounted for only 11 percent.

Table 4: Plan Nacer Expenditures (millions, current USD) 2005–08

Description	2005	2006	2007	2008	Total	Percent
Payments for services	1.89	13.03	24.68	28.40	67.11	63
Investment	3.40	8.72	7.83	2.63	22.63	21
Auditing and IT services	1.90	3.15	5.30	2.86	13.57	12
Training	0.15	0.26	0.31	0.18	0.90	1
Administration	0.20	0. 52	0.60	0.68	1.81	2
Front-end fees	0.68	—	—	—	0.68	1
Total	8.22	25.16	38.72	34.75	106.70	100

Source: Unidad de Financiamiento Internacional de Salud (UFIS), Ministerio de Salud Nacional. 2005-2008.

3. DATA

In this section we describe the construction of the analysis sample and the dependent variables. We focus our analysis on seven of the nine provinces in the northern part of Argentina. The provinces included in the analysis are Catamarca, Corrientes, Formosa, Jujuy, Misiones, Santiago, and Tucumán. As described in the next section, we identify the causal effect of Plan Nacer on health outcomes by exploiting the geographic phasing in of the program over time within each province. We therefore excluded Salta because clinics were incorporated into Plan Nacer over a very short period of time, making our identification strategy inoperative. We also excluded Chaco because data were only available for a two-year period and did not cover the universe of births.

3.1. Sources

In collaboration with the Ministry of Health, we constructed a unique database for analysis that combined information from the following sources:

- Sistema Informatico Perinatal (SIP) — individual birth and medical records for births at maternity hospitals.

- Padron Historico de Beneficiarios de Plan NACER — the historical record of Plan Nacer beneficiary status at the individual level.
- REMEDIAR — individual pharmaceutical records linking an individual's national identify number (Documento Nacional de Identidad — DNI) to the clinic that prescribed/dispensed the medicine.
- Plan Nacer clinic billing and payment administrative records.
- 2001 Population Census — proportion of eligible mothers (without formal health coverage) in the clinic's locality.
- SISA — Sistema Integrado de Información Sanitaria Argentino — the geographic coordinates for those health facilities with complete information.

The prenatal care and birth outcome indicators that are used in the analysis were obtained from SIP birth and medical record files kept at maternity hospitals.[8] Each hospital collects these data in its registries for every birth on an ongoing basis. About half of these records are automated in the SIP, while the rest are on paper. Plan Nacer automated the paper records, and we merged them with records already in the SIP data to obtain a dataset comprising the universe of births in maternity hospitals from 2004 to 2008 in the seven northern provinces.

3.2. Analysis Sample

Our starting point was the universe of births in maternity hospitals. We were able to automate or obtain automated birth records for 78 percent of births in the study period. We then determined whether the mother was a beneficiary of Plan Nacer and the clinic where she obtained her prenatal care. Using the mother's national identity number (DNI) and address we were able to match our existing dataset (described above) to the other administrative databases. Specifically, we constructed the sample used for analysis as follows:

1. *Identify those mothers that became beneficiaries before giving birth.* We used the mother's DNI to check the Plan Nacer beneficiary roster to be able to determine her beneficiary status at the time of pregnancy. A mother was designated to be a beneficiary if she had enrolled in Plan Nacer prior to giving birth, and the clinic that she attended for prenatal care was also enrolled in the Plan at the same time.

2. *Identify the clinic used by the mother for prenatal care.* Birth records were compiled from the hospital where the mother gave birth. However, the clinic

[8] Hospital medical staff collects birth record data, and birth outcome data are not directly incentivized at the provider level. These data are equivalent to birth record data collected in the United States. Thus, any misreporting on a system wide basis on behalf of provinces would involve massive collusion between service providers and the provinces.

where the mother received prenatal care is the relevant service provider for the prenatal care and birth outcomes, since incentives and resources affect prenatal care and birth outcomes by improving clinic-level care. Information on the clinic the mother used for prenatal care was not coded in the birth record. Instead we used a number of databases to assign mothers to clinics. First, we were able to assign Plan Nacer beneficiaries to clinics using Plan Nacer administrative data. For the majority (83 percent) of nonbeneficiaries we were able to identify the clinic through pharmaceutical prescriptions from the REMEDIAR program.[9] For the remainder we used geographic coordinates to identify the clinic closest to where the mother's home. The birth records contained addresses associated with each of the births, and information from SISA provided geographic coordinates for the health facilities.

3. *Identify the timing for incorporation of clinics into Plan Nacer.* While many clinics signed contracts with the Plan early on, most of them did not effectively implement it until much later. Once enrolled, clinics needed time to enroll beneficiaries and learn the program's administrative and billing processes. As the effective start date of the program, we use the date on which a clinic submitted the first bill for payment to plan for services provided to Plan beneficiaries.[10]

We lost about 3 percent of these observations due to missing beneficiary status or because we were unable to identify the clinic where the mother obtained her prenatal care; provinces did not appear to differ with respect to sample attrition. The actual number of observations used in any given analysis may have been smaller due to missing values for a specific outcome.

3.3. Outcome Indicators

We analyze indicators of prenatal care use, prenatal; care quality, and birth outcomes (table 5). The indicators include the number of prenatal care visits, whether a pregnant mother that required a tetanus vaccine received that vaccine, whether birth was by cesarean, birth weight, low birth weight (< 2,500 grams), and neonatal mortality in the hospital. Delivery by cesarean section is an indicator of

[9] We validated the clinic assignments using the REMEDIAR program and the geographic assignment with the beneficiary sample. The beneficiary sample has information from Plan Nacer on the specific clinic used. The clinic assignments using REMEDIAR matched the Plan Nacer administrative data in well over 97 percent of cases.

[10] As clinics usually bill Plan Nacer quarterly to the date of the first bill, we assume that the clinic began Plan-related activities in the quarter prior to the bill.

higher-quality prenatal care. Better prenatal care identifying and preventing delivery risk factors will lower the need for a cesarean section.

Table 5: Means of Outcome Indicators

	Sample size	Overall	2004	2008
Prenatal care				
Number of prenatal care visits	228,656	4.563	4.418	4.853
Tetanus toxoid vaccine	108,535	0.786	0.717	0.694
Cesarean section	282,042	0.235	0.223	0.260
Birth outcomes				
Birth weight (in grams)	274,078	3,277	3,256	3,282
Low birth weight (<2,500 grams)	274,078	0.070	0.075	0.070
APGAR score at 5 minutes > 6	248,840	0.982	0.982	0.983
Mortality				
Neonatal mortality per 1,000 live births	131,943	8.28	9.73	6.77

Source: Authors calculations based on medical records

The sample sizes for tetanus toxoid vaccine are smaller than for the other indicators. This is primarily because the vaccine is prescribed only to a subset of pregnant women.[11] Based on these criteria 44 percent of women in the sample required a tetanus vaccine.

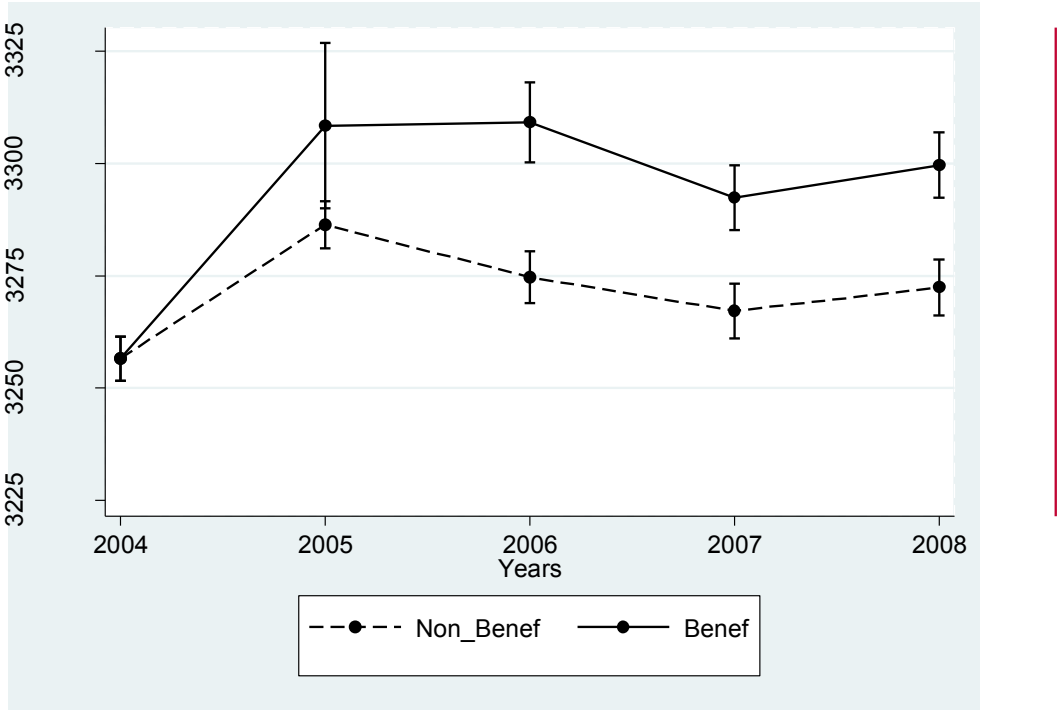
The sample size for neonatal mortality is also lower. We exclude one province and a number of hospitals that coded neonatal mortality aggregated with other outcomes such as abortion. In these cases we could not separate neonatal mortality from abortion. Overall, we have data from 40 hospitals representing around 45 percent of the births. While this should not affect the internal validity of our estimates of the causal impact, it could affect the external validity. However, the neonatal mortality rate from our sample (8.3 per thousand) is close to the neonatal mortality rate based on death certificate data (10.1 per thousand). We would expect the death rate from death certificates to be slightly higher than the in-hospital mortality rate as the former includes deaths that did not occur in hospitals.

[11] Women requiring a tetanus vaccine include those who had not previously been vaccinated or had received only one dose, those whose previous immunization status is not known, those who have had two doses but whose last vaccination is more than three years ago, and those that have had three doses but the last dose or booster was more than five years ago (the maximum length recommended by clinical guidelines).

4. DESCRIPTIVE ANALYSIS

Before turning to results, we examine birth-weight over time, to get a sense of the variation in the data. Figure 5 presents mean birth-weight and the 95 percent confidence region for beneficiaries and non-beneficiaries by year. Over time, as Plan Nacer scales up, the sample sizes of the beneficiary group grow, and the sample size of the non-beneficiary group shrinks. There is no difference between the beneficiary and non-beneficiary means in 2004. Starting in 2005, however, the mean in the beneficiary group is higher than the mean in the nonbeneficiary group, and the gap between beneficiary and nonbeneficiary grows over time. In the last three years (2006–08) the variation in results are significantly different from zero. Although this evidence cannot be interpreted as causal, it is consistent with improvement for the beneficiaries.

Figure 5: Average Birth Weight (*grams*) Over Time



Note: The figure plots mean birth-weight by year and beneficiary status. The individual is defined as beneficiary if she was enrolled in Plan Nacer when she gave birth and if the clinic she visited during pregnancy was incorporated into Plan Nacer.

5. IDENTIFICATION AND ESTIMATION

We investigate three questions:

a. What impact does incorporating a clinic into Plan Nacer have on the outcomes of the population, regardless of beneficiary status?
b. What is the impact of Plan Nacer on the health outcomes of program beneficiaries who receive care from a clinic that is enrolled in Plan Nacer?
c. What is the impact of the program on the health outcomes of non-beneficiaries who receive care from a clinic that is enrolled in Plan Nacer? This question is equivalent to asking whether there are spillovers to non-beneficiaries in treatment clinics.

The first question is an intent-to-treat analysis and requires fewer identification assumptions to estimate causal impacts, whereas the latter two questions are treatment-on-treated analyses that require additional identification assumptions.

5.1. Intent-to-Treat (ITT) Analysis

We estimate the intent-to-treat model using a difference-in-difference approach. The model exploits the geographic phasing of clinics into Plan Nacer over time within each province. Each province was responsible for incorporating clinics and beneficiaries into Plan Nacer. It took the provinces a substantial amount of time to brief, train, and introduces Plan Nacer systems to clinic personnel. In practice, provincial governments typically took several years to complete this process. The treatment groups are the set of clinics that were incorporated early; the control groups are the set of clinics incorporated later.

The difference-in-differences approach compares the change in outcomes in the treatment group to the change in outcomes in a comparison group. The change in the comparison group is an estimate of the true counterfactual — that is, what would have happened to the treatment group if there were no intervention. In our case, the composition of the treatment and control groups changes over time. At baseline in 2004, no clinics were enrolled in Plan Nacer. In the next period, the treatment clinics had converted to Plan Nacer but the control groups had not yet converted. In the subsequent periods, the clinics that had converted to Plan Nacer were added to the treatment group and removed from the control group. Some of the clinics never converted and are controls for the full period of observation. Hence, over time the membership of the treatment and control groups changes.

The difference-in-difference model controls for potential bias from time-invariant characteristics at the clinic level as well as from time-varying characteristics at the provincial level that are common to both treatment and control clinics. This may be important as health services delivery and policy is decentralized to the provincial level in Argentina. Hence, we can control for changes in local health policy and economic growth that might confound the estimates.

We estimate the following difference-in-difference specification:

$$y_{ijkt} = \alpha_j + \gamma_{kt} + \sum_l \delta_l x_{il} + \beta Plan_{jkt} + \varepsilon_{ijkt} \quad , \qquad (1)$$

where

y_{ijkt} is the dependent variable for patient i who was cared for in clinic j living in province k in period t;

α_j is a fixed effect for clinic j;

γ_{kt} is a fixed effect for province k in time period (year) t;

x_{il} is a vector of individual characteristics including maternal age and number of previous pregnancies; [12]

$Plan_{jkt}$ equals 1 if clinic j in province k in period t is incorporated in Plan Nacer and 0 otherwise; and

ε_{it} is a 0 mean random error.

The coefficient β in equation (1) is the difference-in-difference estimate of the impact of Plan Nacer on the outcomes of patients in clinics who convert to Plan Nacer regardless of individual beneficiary status. The clinic fixed effects control for characteristics of a clinic and the population it serves that are constant over time. The province-time fixed effects control for time-varying characteristics common to treatment and control clinics within each province. Maternal age and previous pregnancies control for major risk factors that vary over time.

5.2. Treatment-on- the Treated (TOT) Analysis

We now estimate the impact of Plan Nacer on the outcomes of beneficiaries. We consider first estimating the following specification:

$$y_{ijkt} = \alpha_j + \gamma_{kt} + \sum_l \delta_l x_{il} + \beta Ben_{ijkt} + \varepsilon_{ijkt} \qquad (2)$$

where the treatment indicator Ben_{ijkt} is at the individual level rather than at the clinic level. The coefficient β is now the estimate of the impact on a beneficiary or the treatment-on-treated.

The causal identification β is complicated by the possibility that there could be a health shock that causes an individual to enroll in Plan Nacer and further that the health shock is correlated with prenatal care utilization and birth outcomes. In

[12] Specifically we include maternal age (less than 18 and more than 35 years), first birth, and parity (number of previous births).

this case, the difference-in-difference estimates would be confounded with an idiosyncratic unobserved health shock that is correlated with both treatment status and outcomes. As a result, we would not be able to interpret β as the causal impact of Plan Nacer in beneficiary outcomes.

We can correct for this bias by using a local average treatment effect (LATE) approach. This model assumes that Plan Nacer only affects the outcomes of beneficiaries and has no effect on nonbeneficiaries in treatment clinics. In this case, the ITT effect is the average of the TOT effect on beneficiaries and zero, that is, the treatment effect on nonbeneficiaries in the treatment clinics. Under this assumption, the LATE estimator is simply the ITT estimate divided by the difference in beneficiary enrollment (that is, take-up) between the treatment and control clinics. Since beneficiary enrollment in control clinics is by definition zero, this amounts to dividing the ITT estimate by the enrollment rate in treatment clinics.

We use instrumental variables to estimate the LATE versions of the TOT equation (2), using as the instruments for beneficiary status whether the clinic had already been incorporated into Plan Nacer at the time of the mother's prenatal care and the length of time since the clinic adopted Plan Nacer. Once a provider is enrolled in the Plan, it starts to add its efforts to overall efforts of the province to enroll eligible women and children into Plan Nacer. Since this cannot be done instantaneously, the longer the provider has been on the Plan, the more potential beneficiaries the provider is able to find and enroll, and hence the higher the probability that an individual is enrolled. We verify this statistically significant relationship in the first-stage regressions.

5.3. Spillover

The key assumption in identifying the LATE estimate of TOT is that there is no spillover to nonbeneficiaries in treated clinics. This may not be true and needs to be tested. We can adjust the specification in (2) to allow for spillover to nonbeneficiaries as follows:

$$y_{ijkt} = \alpha_j + \gamma_{kt} + \sum_l \delta_l x_{il} + \beta Ben_{ijkt} + \theta Plan_{jkt} + \varepsilon_{ijkt} \qquad (3)$$

This specification includes indicators for the individual beneficiary, Ben_{ijkt}, and the clinic , $Plan_{jkt}$. The impact on beneficiaries is estimated by $\beta + \theta$, and the impact of Plan Nacer on nonbeneficiaries in treatment clinics is θ. If $\theta = 0$, then there is no spillover to nonbeneficiaries, and the impact on beneficiaries reduces to β.

The estimation of (3) is complicated by the inclusion of both the individual and clinic-level treatment effects since we can no longer use the clinic treatment

indicator as an instrument for individual take-up. To estimate (3), we need to find additional instruments for take-up, that is, factors that affect beneficiary take-up but not outcomes. One possibility is the length of time since the clinic was incorporated into Plan Nacer. Clinics begin the outreach to eligible mothers after the clinic itself is incorporated into Plan Nacer. It can take a long time before clinics are able to complete their outreach, and the longer the length of time, the more likely that an eligible mother enrolls in Plan Nacer. Also, the length of time should be uncorrelated with birth outcomes.

6. ESTIMATION RESULTS

6.1. Intent-to-treat

The intent-to-treat results, presented in panel A of table 6, examine the effect of a clinic enrolling in Plan Nacer on the outcomes of patients regardless of beneficiary status. The results show a significant increase in the number of prenatal care visits and the quality of prenatal care measured by an increase in the share of mothers who receive the tetanus toxoid vaccine and a reduction in the number of births delivered by caesarian. Improved prenatal care appears to be translated into improved birth outcomes as we observe a significant increase in average birth weight and a reduction of the share of low birth weight babies.

6.2. Treatment on the Treated

The treatment-on-treated results, presented in panel B of Table 6, are interpreted as the impact of a clinic adopting Plan Nacer on Plan Nacer beneficiaries assuming that there is no spillover to non-beneficiaries. The results, as expected, show substantially higher impacts. We find that the number of beneficiary prenatal care visits increased significantly by 0.68 visits. The results also show that the probability of a beneficiary receiving a tetanus vaccine increased by 5.6 percentage points, which is a reduction of 24.7 percent in the share of mothers who did not receive a recommended tetanus vaccine. Similarly, we find that a 5.2 percent point reduction the share of beneficiary births delivered by the share of women who had cesarean sections, which is a 21 percent reduction in cesareans. Finally, we find a significant improvement in beneficiary birth weight and reduction in the probability of low birth weight. Specifically, the probability of low birth weight falls by 1.4 percentage points, which translates into a 19 percent reduction overall.

Table 6: Impact of Plan Nacer on Prenatal Care and Birth Outcomes

	Number prenatal care visits	Tetanus toxoid vaccine	Cesarean section	Birth weight (grams)	Low birth weight (<2500 gr)
A. Intent-to-treat (difference-in-difference)					
Clinic on Plan Nacer	0.125**	0.018**	-0.019***	17.11***	-0.005**
	(0.043)	(0.008)	(0.006)	(4.88)	(0.002)
Mean of nontreated	4.486	0.776	0.246	3267.730	0.072
B. Treatment-on-treated (IV difference-in-difference)					
Plan Nacer beneficiary	0.679***	0.056**	-0.052**	48.95***	-0.014**
	(0.151)	(0.027)	(0.022)	(14.78)	(0.006)
Mean of nontreated	4.401	0.774	0.247	3,266.39	0.074
F-statistic for first stage	113.47	132.05	153.53	231.91	231.91
C. Treatment-on-treated with spillover (IV difference-in-difference)					
Plan Nacer beneficiary	2.283**	0.300	0.113	-17.55	0.001
	(1.058)	(0.370)	(0.091)	(63.33)	(.028)
Clinic on Plan Nacer	-0.805**	-0.082	-0.053	22.70	-0.005
	(0.283)	(0.127)	(0.029)	(21.13)	(0.009)
F-statistic for first stage	15.798	2.348	12.437	19.98	19.98
Number observations	228,656	108,535	282,042	274,078	274,078
Number clinics	1,838	1,810	1,898	1,897	1,897

Notes: The table reports the estimated impact of Plan Nacer on outcomes indicated by the columns. Panel A reports the estimated coefficients and standard errors on the clinic treatment indicator from difference-in-difference models specified in equation (1). Panel B reports the estimates for beneficiary status using the clinic status variable and the length (in months) since the clinic was incorporated into Plan Nacer as instruments for beneficiary status in the difference-in-difference model specified in equation (2). Panel C reports the estimated impact of Plan Nacer on beneficiaries and on nonbeneficiaries (clinic on Plan Nacer) using the length (in months) since the clinic was incorporated into Plan Nacer as instrument for beneficiary status in the difference-in-difference model specified in equation (3). All of the models control for clinic fixed effects, time-province fixed effects, maternal age, and number of previous births. Standard errors are clustered at the clinic level.

6.3. Spillover

Finally, in panel C we report the extent of spillovers to nonbeneficiaries in treatment clinics. There are two hypotheses with respect to spillovers. The first is that clinics use Plan Nacer funds to improve quality of care at the facility level and both beneficiaries and non-beneficiaries benefit from the clinic adopting the Plan ("positive" spillovers). Alternatively, Plan Nacer could induce financial incentives for enrolling and treating beneficiaries but not treating non-beneficiaries. Under this hypothesis, providers would increase effort and resources for beneficiaries and reduce effort and resources for treating non-beneficiaries ("negative" spillovers). There does seem to be some evidence of negative spillovers in birthweight or in quality of prenatal care (i.e tetanus and cesaream section). However, we do find negative and statistically significant spillover effect for the number of prenatal care visits.

6.4. Neonatal Mortality

Table 7 reports the estimated impact of Plan Nacer on neonatal mortality. Panel A reports the ITT results and panel B the TOT results. The overall effects of Plan Nacer on neonatal mortality are impressive. The first column (model 1) reports the results for the full sample. The impact of a clinic adopting the program on the neonatal mortality rate regardless of individual beneficiary status is -0.002 or a 22 percent reduction in neonatal mortality. The impact on Plan Nacer beneficiaries is — 0.007 or a 74 percent reduction in neonatal mortality.

Almost all neonatal mortality is concentrated in low-birth-weight babies. The neonatal mortality rate of low birth-weight (LBW) babies is 0.060 while the neonatal mortality rates of non-LBW babies is 0.003. Hence, Plan Nacer can reduce neonatal mortality both by preventing low birth-weight and by increasing survivorship of risky low-birth-weight babies.

To assess the relative importance of the two mechanisms we estimate a second specification in the second column of table 7 (model 2). In this specification we add an indicator for whether the baby was born with low birth weight and the interaction of the LBW indicator and treatment. The coefficient on treatment in both the ITT (panel A) and in the TOT (panel B) models is very small and not statistically significant. This is consistent with the fact that almost all neonatal morality occurs in the population of LBW babies.

In both the ITT and TOT specifications, the coefficient on LBW is positive and significant at 0.062, suggesting that a LBW baby has a 6.2 percent higher chance of neonatal mortality compared to a baby with normal birth-weight. In the ITT

specification, the coefficient for the interaction of Plan Nacer and LBW is -0.010, implying that a LBW baby has a 1 percent lower chance of neonatal mortality if the mother had her prenatal care in a Plan Nacer clinic compared to prenatal care in a non-Plan Nacer clinic. In the TOT model, a LBW baby whose mother had care in a Plan Nacer clinic had less than half the probability of neonatal mortality than a LBW whose mother had care in a non-Plan Nacer facility.

Table 7: Impact of Plan Nacer on Neonatal Mortality

	Model 1	Model 2 LBW
A. Intent-to-treat		
Clinic on Plan Nacer	-0.002** (0.001)	-0.000 (0.001)
Clinic on Plan Nacer * LBW		-0.010** (0.005)
LBW		0.062*** (0.004)
Mean of Dependent Variable	0.0093	0.0093
Nonbeneficary mean/LBW		0.0648
B. Treatment-on-the-treated		
Plan Nacer beneficiary	-0.007** (0.003)	-0.001 (0.003)
Plan Nacer beneficiary * LBW		-0.033*** (0.006)
LBW		0.062*** (0.001)
F-statistic for first stage	127.45	70.02
Sample size	131,943	131,943

Note: Model 1 reports the basic model specifications. In model 2 we add an indicator for low birth weight and its interaction with treatment. Similarly, in model 2 we add an indicator for very low birth weight and its interaction with treatment. Panel A reports the estimated coefficients and standard errors on the clinic treatment indicator from difference-in-difference models specified in equation (1). Panel B reports the estimates for beneficiary status using the clinic status variable and the length (in months) since the clinic was incorporated into Plan Nacer as instruments for beneficiary status in the difference-in-difference model specified in equation (2). All of the models control for clinic fixed effects, time-province fixed effects, maternal age, and number of previous births. Standard errors are clustered at the clinic level.

The effect of Plan Nacer on neonatal mortality can then be decomposed into effects of improved prenatal care that prevent LBW and improved care for those born with LBW. Since neonatal mortality is almost exclusively among LBW babies, we can write the probability of neonatal mortality, $Pr(death)$ as the following:

$$Pr(death) = Pr(death/LBW = 1) \times Pr(LBW) \quad (4)$$

where $Pr(death/LBW = 1)$ is the probability of neonatal morality conditional on LBW, and $Pr(LBW)$ is the probability of LBW.

We then totally differentiate (4) and divide through by the change in the probability of neonatal mortality denoted $\Delta Pr(death)$ to get

$$1 = \frac{\Delta Pr(death|LBW=1) \times Pr(LBW)}{\Delta Pr(death)} + \frac{\Delta Pr(LBW) \times Pr(death|LBW=1)}{\Delta Pr(death)} \quad (5)$$

The first term on the right side of equation (5) is the share of the reduction in neonatal mortality caused by improved care for LBW babies, and the second term is the share caused by reducing the probability of a LBW birth.

The elements of (5) can be read from the tables. The denominator is estimated by the treatment effect in model 1 in table 7. The numerator of the first term is the interaction of treatment and LBW in model (2) times the share of babies born with LBW. The second denominator is the treatment effect from table 7 times the neonatal mortality rate for LBW babies. In the ITT models, these estimates are the impact of Plan Nacer on the neonatal mortality of non-LBW babies. The decomposition results imply that 54 percent of the reduction in neonatal mortality was caused by a reduction in the probability of LBW; the remaining 46 percent was caused by improved care for babies that were born with LBW.

7. IDENTIFICATION ROBUSTNESS TESTS

7.1. Assumptions

Both the difference-in-difference models and the treatment-on-treated with and without spillover require assumptions to interpret the results as causal. The difference-in-difference model uses the change in the outcomes of the control group to estimate the counterfactual, that is, what would have been the change in the outcomes of the treatment group had the clinic not adopted Plan Nacer. Hence, the most important identifying assumption is that the change in the outcomes of the control groups is a consistent estimate of what would have been the change in the outcomes of the treatment group had the clinics not adopted Plan Nacer.

Similarly, we use whether the clinic had already adopted Plan Nacer and the length of time in Plan Nacer as instruments for beneficiary take-up in both the TOT

model and the spillover model. The instruments must be correlated with take-up and otherwise uncorrelated with the outcomes, conditional on the clinic and year-province fixed effects. Since beneficiaries cannot enroll if the clinic they use is not a part of Plan Nacer, and it takes time for the clinic to enroll beneficiaries, we expect that the fact the clinic is a member of Plan Nacer and the time it has been on Plan Nacer to predict individual take-up. Moreover, each province developed plans to enroll clinics into Plan Nacer over time, and those plans were primarily based on cost and convenience. Expansion plans did not explicitly target early enrollment of areas in which birth outcomes were worse or better or areas in which outcomes were trending worse or better.

We can obviously directly test the first assumption that the instruments predict take-up. We can partially test the second assumption by investigating whether clinic take-up was correlated with baseline birth outcomes or the trends in those outcomes. If take-up is not correlated with baseline birth outcomes or trends in outcomes, it would confirm the institutional information provided by the provinces that they did not plan the rollout baseline on levels or trends in outcomes. Since the analyses are conditional on clinic fixed effects, it would be enough for the timing of future take-up to be uncorrelated with pre-intervention trends in outcomes. The same requirement exists for the difference-in-difference models. However, we cannot test whether actual differential idiosyncratic time-varying shocks to local birth outcomes at the clinic level affect take-up. This is unlikely to be the case as the provinces developed the clinic rollout plan, and those plans were implemented over time. There is no evidence that any province changed those plans in response to unexpected shocks to local birth outcomes.

7.2. Differential Pre-intervention Trends Specification Tests

While we cannot directly test difference in the trends after intervention, we test equality of treatment and control pre-intervention trends in the outcome variables. If there is no difference in the pre-intervention trends of the treatment and control groups, then there is little reason to expect a difference in trends in the post-intervention period absent the intervention.

We implement this test by estimating the following equation using only pre-intervention observations (that is, before the clinic is incorporated into Plan Nacer):

$$y_{ijkt} = \alpha_j + \gamma_{kt} + \sum_l \delta_l x_{il} + \sum_{s=2005}^{2008} \beta_{ks} Trend_t Plan_{jks} + \varepsilon_{ijkt} \qquad (6)$$

Equation (6) replaces βT_{jkt} with $\sum_{s=2005}^{2008} \beta_{ks} Trend_t Plan_{jks}$ where $Trend_t$ is a quarterly time trend variable and $Plan_{jks}$ is an indicator of whether clinic j in province k will be incorporated into Plan Nacer in year s. These terms allow time

trend differences for clinics that become treated in 2005, 2006, 2007, and 2008. The hypothesis of no differential trends is tested by seeing if the coefficients on the βs are jointly not different from zero.

The results of these tests are presented in panel A of annex tables A.1 for the prenatal care variables, birth outcomes, and neonatal mortality. The table presents the *p*-values for the tests of the hypotheses that the trends in outcomes differ depending on if and when a clinic is incorporated into Plan Nacer. For each outcome the table reports the *p*-value of the test of whether the trend in outcomes of mothers who obtained their prenatal care at clinics that were incorporated in Plan Nacer in a particular year were different than the trends in outcomes of mothers who received their care from other clinics and for a joint test of no difference for all years of incorporation. Only 2 of the 30 p-values reject the null hypothesis of differential pre-intervention trends at the 0.10 level. This provides strong evidence of no difference in pre-intervention trends and provides strong support to the assumptions necessary to interpret both the difference-in-difference and treatment-on-treated results as causal.

7.3. Baseline Balance

We can also check baseline balance of the treatment and control groups. While baseline balance is desirable, it is not necessary for the difference-in-difference results to be interpreted as causal. The provider fixed effects account for differential initial conditions. We check for balance by estimating the following equation using the pure baseline, that is, the observations before any clinic was incorporated into Plan Nacer in 2004:

$$y_{ijk,2004} = \alpha_j + \sum_l \delta_l x_{il} + \sum_{s=2005}^{2008} \beta_{ks} Plan_{jks} + \varepsilon_{ijk,2004} \qquad (7)$$

Again, the hypothesis of baseline balance is not rejected if the βs are jointly zero.

The results of these tests are presented in panel B of annex tables A.1 the prenatal care variables, birth outcomes, and neonatal mortality, respectively. The table presents the *p*-values for the tests of the hypothesis that baseline outcomes differ depending on if and when a clinic is incorporated into Plan Nacer. For each of the outcomes the table reports a *p*-value of the test of whether the baseline outcomes of mothers who obtained their prenatal care at clinics that were incorporated in Plan Nacer in a particular year were different than the baseline outcomes of mothers who received their care from other clinics, and for a joint test of no difference for all years of incorporation. In only 1 out of the 30 tests, was baseline balance rejected. This provides strong evidence of no difference in baseline

means and provides further support to the assumptions necessary to interpret both the difference-in-difference and treatment-on-treated results as causal.

8. COST-EFFECTIVENESS

In order to assess the policy relevance of the estimated effects, we estimate the cost-effectiveness of Plan Nacer in terms of reducing low birth weight and neonatal mortality. One issue is that we have two outcomes that have different metrics. We follow the convention in the public health literature of converting the estimates of program impact into disability-adjusted life years (DALYs) saved (Jamison, 1993; Murray, C. J. L. 1994; Murray and Lopez 1996a-c). We then estimate cost-effectiveness by dividing DALYs averted due to Plan Nacer by the incremental costs of the program. As recommended by the World Health Organization based on the report of the Commission on Macroeconomics and Health (WHO 2001), we benchmark the cost per DALY against national per capita GDP to assess whether the intervention is cost-effective in the Argentine context. We begin by estimating incremental cost and then DALYs saved.

8.1. Costs

In this section we describe how we estimate Plan Nacer costs for maternal services, that is, prenatal care, plus delivery, and hospital-based neonatal services in the seven provinces under study. We consider both fixed and variable (recurrent) costs. Table 8 presents our estimates of the costs computed using information provided by the Ministry of Health in local currency. We first converted them into US dollars using the official exchange rate and then deflated them to 2005 using the national GDP deflator.

Table 8: Plan Nacer Costs for Maternal Health Services in Seven Provinces (Millions, 2005 US$)

	2005	2006	2007	2008	Total
Recurrent costs					
Payments for services	0.46	3.33	6.60	7.37	**17.77**
Other joint recurrent costs	0.18	0.50	1.16	0.69	**2.27**
Fixed costs					
Use cost of durable goods	0.010	0.100	0.030	0.013	**0.150**
Total costs (USD)	**0.87**	**3.93**	**7.79**	**8.08**	**20.45**

Source: Unidad de Financiamiento Internacional de Salud (UFIS), Ministerio de Salud Nacional.

Recurrent costs include payments to facilities for prenatal care, delivery and hospital postnatal care services as well as other joint recurrent costs such as external auditing, IT services, and management personnel. These other joint recurrent costs are only available at the aggregate (national) level year. We estimate other joint recurrent costs for maternal services by multiplying total joint recurrent costs by maternal services share of total billings for services each year for the seven provinces.

Fixed costs include investment in durable goods such as medical equipment, information technology, vehicles, and office equipment. We calculate *E*, the "annual use cost" of a durable, using the following formula:

$$E = \frac{K - \left[S/(1+r)^n\right]}{A(n,r)}$$

where *K* is the purchase price, *n* is the useful life of the durable, *S* is the end of life resale value, *r* is the discount rate, and *A(n, r)* is the annuity factor and given by the equation: $\left[\frac{1-(1+r)^{-n}}{r}\right]$. We followed WHO guidelines to estimate the useful life and resale value of an asset (*n* and *S*). We classified durables into five groups: medical equipment (10 years), medical equipment of high complexity (15 years), ambulances (15 years), information system equipment (5 years), and others (small items/program credentials). We assume a zero-resale value at the end of the project and applied a 3 percent discount rate.

We were able classify medical equipment according to its medical use — maternal care, baby or children care, or all care. However, information technology, vehicles, and office equipment were only available at the aggregate level. Again, we estimated fixed costs from theses associated with maternal health services by multiplying total costs by maternal services share of total billings for services each year for the seven provinces.

8.2. Disability-Adjusted Life Years Averted

DALYs are the sum of the present value of future years of lifetime lost through premature mortality, and the present value of future years of lifetime living with a disability caused by disease or injury adjusted for the average severity of the disability. DALYs allow us to sum years of life lost from neonatal mortality with the years living with disability associated with low birth weight.

We calculate DALYs averted due to Plan Nacer as:

$$DALY = ND * \sum_{t=0}^{LE} \left(\frac{1}{1+r}\right) + LBW * D_{lbw} * \sum_{t=0}^{LE} \left(\frac{1}{1+r}\right) \quad (8)$$

where ND = number of neonatal deaths averted by Plan Nacer

LBW = number of LBW babies averted by Plan Nacer

r = discount rate

LE = Life expectancy at birth

D_{lbw} = Disability weight for low birth weight

The disability weight for low birth weight is an estimate of the loss in value of life from living with the disability associated with low birth weight and provides a common metric for summing the loss due to disability and mortality across disease and injury types.

The DALYs associated with a neonatal death are the present value of life expectancy at birth. Life expectancy at birth in Argentina is 76 years and assuming a 3 percent discount rate, the present value of years of life lost from a neonatal mortality is 30. A child who experiences low birth weight will live with the associated disability for her whole life. The discounted present value of the years of life living with disability associated with is 30 as well. The disability weight associated with low birth weight is estimated as 0.106 (WHO 2004). Then the DALYs associated with a low birth weight birth in Argentina are 3.18.

The number of neonatal deaths averted is estimated as the product of the total number of births from beneficiary mothers and the absolute value of the estimated coefficient on the effect of Plan Nacer on the probability of an in-hospital neonatal death. Between 2005 and 2008 there were 104,748 beneficiary mothers of Plan Nacer giving birth in the seven provinces under analysis.[13] Hence we estimate that Plan Nacer averted 773 neonate deaths with a 95% confidence region of 117 to 1,349. Then the DALYs associated with neonate deaths averters are 21,997.

The number of LBW babies is estimated as the product of the total number of births from beneficiary mothers and the absolute value of the estimated coefficient on the effect of Plan Nacer on the probability of a low weight birth from Table 8. Recall that half of the averted neonate mortality was through a reduction in low birth weight babies. Hence we need to reduce the estimated number of LBW babies averted by the number of neonate deaths averted due to preventing LBW to avoid double counting. We estimate that the number of LBW babies averted adjusted for neonatal mortality to be 1,071 with a 95% confidence region 171 - 1,970. Then the total number DALYs averted through preventing LBW is 3,404.

[13] See Informes de Gestion, 2005-2008, total beneficiaries by December each year. Total averted deaths are calculated as 733 = 104,748*0.007 for the point estimate.

8.3. Cost-Effectiveness Estimates

We estimate the cost of a DALY saved through Plan Nacer's financing of maternal health services was $814 with a 95 percent confidence region of $442 to $5,086 (table 9). As a way to assess the effectiveness of this program, we follow the World Health Organization recommendation based on the Commission of Macroeconomics and health: investing in health for economic development (2001, Geneva). These results are hugely cost-effective in the Argentine context when compared to GDP per capita over this period. The cost per DALY saved is only 18 percent of average income.

Table 9: **Plan Nacer Maternal Health Services Cost-Effectiveness 2005–08**

	Lower bound	Point estimate	Upper bound
Neonatal deaths averted	117	733	1,349
Low-birth weight babies averted	171	1,071	1,970
Plan Nacer DALYs saved	4,064	25,401	46,738
Cost per DALY averted	$5,086	$814	$442
Annual real GDP per capita (2005–08)	$6,075	$6,075	$6,075

9. SUMMARY AND CONCLUSIONS

In this study, we analyzed the impact of Plan Nacer on birth outcomes using data from the universe of birth records in seven northern Argentine provinces for the period 2004-08. We exploited the geographic phasing in of Plan Nacer at the provincial level to identify plausible causal impacts. We investigated three related questions: (1) What is the impact of incorporating a clinic into Plan Nacer on the outcomes of the population regardless of beneficiary status? (2) What is the impact of Plan Nacer on the health outcomes of program beneficiaries who receive care from a clinic that is enrolled in Plan Nacer? and (3) What is the spillover impact of the program on the health outcomes of nonbeneficiaries who receive care from a clinic that is enrolled in Plan Nacer?

Overall, we conclude that the Plan Nacer incentive-based model has had large positive effects on birth outcomes and is a promising model for emulation both within the Argentine health sector and internationally. We find that the program increases the use and quality of prenatal care services as measured by the number of prenatal care visits and the probability of receiving a tetanus vaccine (as a

measure of quality of care during those visits). We also find substantial improvement in birth outcomes. Specifically, we estimate that being a beneficiary reduces the probability of low birth weight by 19 percent and in-hospital neonatal mortality by 74 percent. About half the reduction in neonatal mortality comes from better prenatal care that prevents low birth weight and half from better postnatal care available to low-birth weight babies. Our results also suggest that there are no negative spillovers onto the birth outcomes among the nonbeneficiary population receiving care in clinics covered by Plan Nacer.

Overall, we find that Plan Nacer is hugely cost-effective. The cost of a DALY saved through Plan Nacer's financing of maternal health services was $814, which is hugely cost-effective in the Argentine context when compared to GDP per capita, $6,075, over this period. Plan Nacer uses a relatively small amount of resources (2 to 4 percent of total expenditure) to provide incentives to health providers to use resources more efficiently and for higher-quality care to program beneficiaries.

There are, however, a number of limitations to our analysis. First, while we have high-quality data on birth outcomes, we have limited information on the use and quality of services that act as mediators between Plan Nacer and ultimate health impacts. Second, the neonatal mortality data are for a subset of the population. While the causal impacts are internally valid, they may not be completely externally valid. However, the neonatal mortality rates for the subsample used in our analyses are not different from the neonatal mortality rates for the region overall.

Annex Table A.1

***P*-Values for Tests of Baseline Balance and Differential Pre-intervention Trends**

	Number of prenatal care visits	Tetanus toxoid vaccine	Cesarean section	Birth weight (grams)	Low birth weight (< 2500 gm)	Neonatal mortality
Panel A: Differential Preintervention Trends Tests						
2005	0.612	0.318	0.229	0.439	0.144	0.767
2006	0.939	0.499	0.313	0.804	0.745	0.340
2007	**0.044**	0.438	0.156	0.582	0.964	0.502
2008	0.131	0.565	0.930	0.583	0.894	**0.065**
All	0.134	0.686	0.260	0.819	0. 597	0.256
N	87,341	44,819	130,343	114,973	114,973	65,953
Panel B: Baseline Balance Tests						
2005	0.264	0.109	0.611	0.406	0.103	0.106
2006	0.045	0.400	0.711	0.731	0.810	0.821
2007	0.820	0.792	0.756	0.399	0.865	0.230
2008	0.177	0.152	0.710	0.793	0.356	0.002
All	**0.060**	0.150	0. 988	0.515	0.276	0.000
N	40,421	16,077	62,354	53,830	53,830	27,121

Note: The table reports the *p*-values for tests of baseline balance and differential pre-intervention trends for the outcomes indicated by the columns. Panel A reports *p*-values for tests of baseline balance based on equation (6). The *p*-values for the years are associated with hypotheses that the pre-intervention means in 2004 are no different for those births whose mothers obtained their care in clinics incorporated into Plan Nacer in that year compared to births whose mothers obtained their care in other clinics. The *p*-value associated with "all" is for the joint hypothesis that there is no difference in baseline means regardless of when the clinic was incorporated into Plan Nacer. Panel B reports the same *p*-values for the tests of differential pre-intervention trends based on equation (7). All *p*-values were computed clustering at the clinic level.

References

Almond, Douglas, Kenneth Y. Chay, and David S. Lee. 2005. "The Costs of Low Birth Weight." *Quarterly Journal of Economics* 120 (3): 1031–83.

Argentina. National Institute of Statistics and Censuses (Instituto Nacional de Estadística y Censos — INDEC). 2010. "National Census 2010." Buenos Aires.

Argentina. National Ministry of Health. 2008a Secretary of Health Promotion and Programs. Plan Nacer Central Implementation Unit (Unidad Ejecutora Central [UEC]).. "Revisión de Estimaciones de Población Objetivo PISMIP 2008 del SDT 23." Technical Document 28. Buenos Aires.

———. "Informes de Gestión: Plan Nacer —Years 2005 to 2009." Buenos Aires.

Argentina. National Ministry of Health. 2008b Plan Nacer Program. *Plan Nacer Operation Manual.* Buenos Aires.

Argentina. National Ministry of Health. 2008c. Plan Nacer Program. "Cambio en el Modelo de Gestión de Trazadoras." Technical Document 31. Buenos Aires.

Argentina. National Ministry of Health. Programa de Salud Materno Infantil Provincial. 2005. "Manual Operativo. Plan Nacer Proyecto de Inversión en Salud Materno Infantil Provincial." Buenos Aires.

Basinga, Paulin, Paul Gertler, Agnes Binagwaho, Agnes Soucat, Jennifer Sturdy, and Christel Vermeersch. 2011, "Effect on Maternal and Child Health Services in Rwanda of Payment to Primary Health-Care Providers for Performance: An Impact Evaluation," *The Lancet* 377: 1421–28.

Behrman, Jere R., and Mark R. Rosenzweig. 2004. "Returns to Birth weight." *Review of Economics and Statistics* 86 (2): 586–601.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 2007. "From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes." *Quarterly Journal of Economics* 122 (1): 409–39.

Chaudhury, Nazmul, Jeffrey Hammer, Michael Kremer, Karthik Muralidharan, and F. Halsey Rogers. 2006. "Missing in Action: Teacher and Health Worker Absence in Developing Countries." *The Journal of Economic Perspectives* 20 (1): 91–116.

Currie, Janet. 2011. "Inequality at Birth: Some Causes and Consequences." *American Economic Review*, Ely lecture, May 2011

Das, Jishnu, Jeffrey Hammer, and Kenneth Leonard. 2008. "The Quality of Medical Advice in Low-Income Countries." *Journal of Economic Perspectives* 22 (2): 93–114.

Dirección Estadistícas e Información de Salud (DEIS). "Indicadores Basicos" (2006–10). National Ministry of Health and Panamerican Health Organization.

Gertler and Veermeersch. 2013. “Using Performance Incentives to Improve Medical Care Productivity and Health Outcomes.” National Bureau of Economic Research Working Paper Series 16798.

Jamison 1993

Miller, Grant, and Kimberly Singer Babiarz. 2013. “Pay-for-Performance Incentives in Low- and Middle-Income Country Health Programs.” NBER Working Paper 18932. April.

Miller, Grant, Renfu Luo, Linxiu Zhang, Sean Sylvia, Yaojiang Shi, Patrica Foo, Qiran Zhao, Reynaldo Martorell, Alexis Medina, and Scott Rozelle. 2012. “Effectiveness of Provider Incentives for Anaemia Reduction in Rural China: A Cluster Randomised Trial.” *British Medical Journal* 345.

Murray, C. J. L. 1994. “Quantifying the Burden of Disease: The Technical Basis for Disability-Adjusted Life Years.” *Bulletin of the World Health Organization* 72: 429–45.

Muralidharan, Karthik, and Venkatesh Sundararaman. 2011. "Teacher Performance Pay: Experimental Evidence from India." *Journal of Political Economy* 119 (1): 39–77.

Musgrove, Phil. 2010. “Plan Nacer, Argentina: Provincial Maternal and Child Health Insurance Using Results-Based Financing (RBF).” mimeo.

Olken, Benjamin, Junko Onishi, and Susan Wong. 2013. “Should Aid Reward Performance? Evidence from a Field Experiment on Health and Education in Indonesia,” Working paper, MIT, Department of Economics, Cambridge, MA.

Spears, Dean. 2013. "Effects of Rural Sanitation on Infant Mortality and Human Capital: Evidence from a Local Governance Incentive in India." Department of Economics, Princeton University, Department of Economics, Princeton, NJ.

World Bank (2004). World Development Report 2004. Making Services Work for Poor People. Washington DC

World Health Organization 2001. Macroeconomics and health: investing in health for economic development. Geneva.

World Health Organization 2004. Global Burden Of Disease Update: Disability Weights For Diseases And Conditions, Geneva http://www.who.int/healthinfo/global_burden_disease/GBD2004_DisabilityWeights.pdf