Scaling Up Oportunidades and Its Impact on Child Nutrition

Gabriela Farfán
María Eugenia Genoni
Luis Rubalcava
Graciela Teruel
Duncan Thomas
Abstract

Oportunidades was an innovative anti-poverty program that put additional resources in the hands of women and their families and encouraged parents to invest in the human capital of their children. This program was the first in its kind, and early evaluations demonstrating its success informed its large expansion within Mexico and the implementation of similar conditional cash transfer programs across the world. However, the existing evidence, which arguably captures causal positive impacts, relies on a sample of very poor rural children. This paper conducts the first evaluation of the program using representative data. It focuses on child height as a marker of long-term nutritional status. The causal impact of the program on child height is isolated by exploiting insights from the biology of child growth in combination with the timing of the rollout of Oportunidades and the panel dimension of the survey. Height for age among children exposed during the first four years of life is contrasted with similar children who were not exposed. Consistent with previous evidence, this analysis finds positive and sizable effects on children who lived in rural poor communities incorporated at the beginning of the intervention. In contrast, the impacts of the program in rural localities incorporated later and in suburban and urban communities are, at best, very modest.
Scaling Up *Oportunidades* and Its Impact on Child Nutrition

Gabriela Farfán†
World Bank

María Eugenia Genoni
World Bank

Luis Rubalcava
CAMBS and Spectron

Graciela Teruel
Universidad Iberoamericana

Duncan Thomas
Duke University

JEL: I1, I3, I38, I18

Keywords: Cash transfers, Nutrition programs, Nutrition policy and programming

† Financial support from the Hewlett Foundation and Institute of International Education (IIE) is gratefully acknowledged. The authors are grateful for comments provided by María Laura Alzúa, Peter Arcidiacono, Federico Bugni, Guillermo Cruces, Erica Field, Susan Parker, Alessandro Tarozzi, Rebeca Thornton, and participants in the Duke Development Reading Group, LACEA 2012, Impact Evaluation Network 2012, NEUDC 2011, PAA 2011, and X Reunión Nacional de Investigación Demográfica, Mexico DF 2010. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.
1. Introduction

Oportunidades (later known as PROSPERA) was an ambitious antipoverty program that started operating in Mexico in 1997. This intervention was the first of its kind, linking cash transfers with investments in human capital as a way to integrate both short-term and long-term anti-poverty policies. Cash transfers were given to satisfy current consumption needs, while investments in education, health and nutrition aimed to establish the grounds for the program to have long-lasting effects. The first phase of the implementation was accompanied by an evaluation design to assess the impact of the program and inform policy. The success in short-term evaluations of this program led to a massive expansion throughout the country, covering about a quarter of the Mexican population. Moreover, this evidence supported a new trend in the design of anti-poverty programs throughout the world. Just ten years later, more than 30 countries were implementing similar conditional cash transfer (CCT) programs.

The expansion of Oportunidades was supported by the assumption that the positive impacts found on communities incorporated at the beginning of the intervention would carry over to the rest of the country. However, the program followed a very specific geographic targeting rule, covering more disadvantaged rural places earlier. Importantly, the most rigorous evidence on program effects comes from evaluations of those early rural areas.

This suggests there is still a very important question unanswered in the literature, which is to see whether the program was successful in improving the well-being of beneficiary households once the program expanded to cover better-off communities and urban areas. This is especially important, considering we are evaluating a very expensive intervention. This research contributes
to an emerging literature that explores the potential threats of scaling up interventions (Araujo, Rubio-Codina, and Shady, 2021; Al- Ubaydli et al., 2019).

The purpose of this work is to fill this gap in the literature for the case of child nutrition. While its impact is manifested in the short-run, an improvement in the nutritional status of children has long-lasting effects. Evidence shows that malnutrition in early childhood is associated with deficits in cognitive development, greater risk of infant and child mortality and morbidity, as well as worse health status and lower earnings during adulthood (Martorell, 1999; Martorell et al., 2005; Strauss and Thomas, 1995).

There are a few papers in the literature that evaluate the impact of the program on child nutrition, and focus on height as one of their outcomes of interest. However, all the available evidence is based on the Oportunidades evaluation data which only represents the poorest communities in Mexico. These data consist of two samples. A rural evaluation sample selected in 1997, at the time the program first started; and an urban evaluation sample, selected in 2001 at the time the program expanded to cover urban areas. Both baseline samples are complemented with a series of follow-up surveys. The expansion of the program followed an explicit geographic targeting rule, where more disadvantaged areas were covered earlier. Thus, a key aspect of these data is that they represent very specific geographic areas of Mexico. As a result, the current evidence only speaks to the impact of the program on communities incorporated early in time, but there is no evidence on the performance of the intervention in places incorporated later, which are relatively better-off.

In contrast to previous evidence, this is the first time population-representative data are used to analyze this program, which allows performing an impact analysis at the national level. The data

---

1 See Parker et al. (2008) for a summary of the literature.
used in this paper is the Mexican Family Life Survey (MxFLS), a longitudinal survey that collects an extensive set of information on individuals, households and communities. We use baseline information collected in 2002, and the first follow-up carried out in 2005-2006. Important for our project, these data are, at baseline, representative at the national, rural-urban and regional level. We complement these household data with Oportunidades administrative records to identify in which year each of our MxFLS localities was incorporated into the program.

In order to isolate impact effects using non-experimental data we propose an identification strategy that combines insights from the biology of child growth, the timing in the roll-out of the program, and the panel dimension of our household survey. The nutritional literature has established that height is responsive to nutritional interventions early in life, but is largely determined once children reach the age of 4. Drawing from this evidence, we define exposure to the program as a function of the age of the child at the time the program arrives to the locality of residence, and exploit the variation in exposure generated by age differences. Next, we divide our localities into different groups based on their date of entry into the program. This allows to control for program placement effects and time trends, and to link our results to the existing evidence. Finally, we use the fact that we measure children at two points in time to control for the nonlinearity of height over the life course. This identification strategy is innovative within the Oportunidades literature, but the basic intuition behind it has been successfully implemented in the nutrition and economics literatures (Martorell and Habicht, 1986; Duflo, 2001; Frankenberg et al., 2005). Importantly, it allows to compare impacts as the program expanded.

The results presented in this paper suggest that Oportunidades did have a positive impact on young children that lived in the poorest communities in rural Mexico. These results match existing evidence that exploits the experimental design of the program during the first couple of years of
operation. With respect to the communities in the rural areas that were incorporated in later stages, estimated effects suggest that the program still had a positive impact in communities incorporated between 1999 and 2002. In contrast, there is no evidence of a positive impact of the intervention on children that lived in areas incorporated much later, between 2003 and 2005, more than 6 years after Oportunidades started. Finally, there is very limited evidence of an improvement in children’s nutritional status in urban areas. Estimated effects are never significant, even when we restrict the sample to the more disadvantaged households.

The remaining of the paper is organized as follows: Section 2 provides a brief description of the program. Section 3 introduces the data used and Section 4 describes the identification strategy. Next, Section 5 discusses some caveats and robustness checks. Sections 6 and 7 present and discuss the results. Section 8 concludes.

### 2. Program background

The Oportunidades program (previously known as PROGRESA) was first implemented in 1997 in the poorest rural areas of Mexico, and since then it gradually expanded to cover less marginal rural areas and urban areas. By the end of 1999 about 40 percent of the rural population, approximately 2.6 million families in almost 50,000 localities, were enrolled in the program. By the end of 2002 the program was operating in 70,520 localities, in all 31 states of the country, reaching 4.24 million households. As of mid-2005, Oportunidades covered 5 million families - about one-quarter of the Mexican population.²

² For information on the characteristics of the program, see Angelucci and Attanasio (2009), Behrman and Todd (1999), Gutierrez et al. (2003), Rivera et al. (2000), Skoufias (2005), Skoufias et al. (1999a), Skoufias et al. (1999b).
Oportunidades was the first CCT program to be implemented. By design, it linked cash transfers to investments in human capital with the rationale that such integration addresses simultaneously short term and long term constraints to break the intergenerational transmission of poverty. The original design, which was in place during the period over which we evaluate the intervention, had three components.

First, a food component consisted of a fixed amount of money given to every beneficiary household regardless of its demographic composition, and it aimed to improve the food consumption and nutritional status of poor families. This transfer was conditional on following the health and nutrition investments specified under the health and nutrition component.

Second, an education component consisted of a pre-specified amount of money that households received for each child enrolled in grades 3 to 12. The amount varied with the gender of the child and the school grade, and there was a maximum amount a household could get. These transfers were conditional on regular attendance (minimum of 85 percent) of school-age children.³

Third, a health and nutrition component consisted of access to basic health care services, provision of nutritional supplements, and educational talks. The nutritional supplements were given to pregnant and lactating women and children between 4 and 24 months. They were also given to children between 2 and 4 years old if malnutrition symptoms were detected by health personnel. The educational talks were community meetings where trained nurses and physicians discussed topics related to health, hygiene, and nutrition issues and practices. Program participation required every household member to go through regular health check-ups, women to have regular prenatal

³ In 2003 Oportunidades added the component “youth with opportunities”. It consists of a savings account that can be cashed when students graduate from high-school (12 years of education) if they graduate before they turn 22 years old.
and post-pregnancy visits, the growth monitoring of preschool children, and regular attendance to the educational meetings. The required periodicity of health check-ups and attendance to educational meetings varied by household member.

Importantly, the roll-out of the program followed a very specific geographic targeting rule, reaching more disadvantaged communities earlier and gradually incorporating better-off communities over time. At first, only rural localities were eligible for the intervention. This restriction was later relaxed, with semi urban areas incorporated since 2001 and urban areas incorporated since 2002.

Among rural places, the selection of eligible localities was based on a marginality index computed from census data. In contrast, due to the high degree of heterogeneity in socioeconomic status within urban localities, the geographic unit of interest to select urban places was the Primary Sampling Unit (PSU), where a PSU consisted of a group of 1 to 50 blocks. Only localities with at least one PSU with high concentration of poor households were selected.

After a community was targeted, eligible households were selected. In eligible rural localities the program carried out a census, computed a household-level poverty index, identified those households that would be eligible to enroll, and then informed them about their eligibility status. As a result, enrollment was almost automatic and administrative records suggest that 97 percent of eligible households were incorporated to the program. Facing a much higher cost-benefit ratio in conducting a census in urban localities, the program changed the enrollment process in 2002. Instead of collecting information on every household, the program established registration offices

---

4 The program used information collected in the 1990 Mexican Population Census and the 1995 Population Count.
5 Semi urban areas incorporated in 2001 followed the same selection process as in rural areas. Similarly, the household selection process applied to rural areas incorporated to Oportunidades in 2002 or later was a variant of that applied in urban areas (Gutierrez et al., 2003).
in eligible areas and advertised the program through campaigns. Interested households had to go
to these offices and provide the information to determine eligibility, a costly process that resulted
in much lower take-up rates and led to significant self-selection in program participation among potential beneficiary households. Administrative data suggests that, initially, only about 50 percent
of eligible households were enrolled in the program.⁶

3. Data

The main data source of this paper is the Mexican Family Life Survey (MxFLS), a longitudinal
survey that collects a rich set of information on demographic and socioeconomic characteristics of individuals, households, and communities. At baseline, the sample is representative at the national, rural-urban and regional level. The first wave (MxFLS1) was conducted in 2002 and interviewed
35,677 individuals in 8,440 households. These households reside in a total of 150 communities
located across 16 different states.⁷ The second wave (MxFLS2) was conducted during 2005-2006
and achieved a 90 percent re-contact rate at the household level. This wave consists of 36,946
individuals in 8,434 households, who due to migration decisions are located across 247 localities
in 21 states throughout Mexico. A third wave (MxFLS3) spanned through 2009-2013. Due to the
timing in the roll-out of Oportunidades, this paper draws on data collected in the first two rounds
of the survey.

The two key pieces of information in our analysis are: height – our outcome of interest, and the
timing in the roll-out of the program in MxFLS localities. With respect to height, MxFLS collects

⁶ More specifically, households interested in the program had to go to the registration offices on specific dates and
answer an inclusion questionnaire. With that information households were immediately classified as qualified for the
program or not. If qualified, they had to answer a second questionnaire and were visited later in their dwellings to
confirm their eligibility status.
⁷ Mexico is divided into 31 states and the Federal District.
high-quality anthropometric measures on every household member, which were taken by trained personnel from the National Institute of Public Health (INSP). In order to control for age and gender specific differences in child height, height-for-age z-scores are computed using the 2000 CDC Growth Charts for the United States provided by the National Center for Health Statistics (NCHS). To compute those z-scores, we calculate age in months using the information on date of birth and interview date available in the data.

With respect to the roll-out of the program, the MxFLS survey collects information on Oportunidades current participation at the individual, household and community level. However, we do not know participation histories. Therefore, in order to identify the year in which each MxFLS community was incorporated to the program, we combine MxFLS data with Oportunidades administrative records to assign the year of entry of the program to a locality. In particular, we use two sources of information. One source is the complete list of Oportunidades beneficiaries as of December 2009. These records have individual-level information on locality of residence and date of enrollment in the program. Based on the households’ date of entry, each of the 246 MxFLS localities is associated with the year in which the largest number of households was enrolled in the program. The second source of information is a locality-level data set with the number of families enrolled each year. Combining these two data sets, together with the self-reported participation rates reported in MxFLS, we assign a year of incorporation to each MxFLS locality. Figure 1 illustrates the pattern of expansion of the program for the rural and urban areas separately. MxFLS self-reported participation rates in 2002 and 2005 are very similar to the ones obtained from administrative records, which should be expected as we have a population-representative sample in 2002.
Table 1 characterizes the final sample. As will be explained later, our analysis focuses on children ages 1-3 and 5-7. The original sample of children in this age range in rural communities is 3,500 and in urban communities is 4,849. Some cases are dropped from the analysis due to lack of measurement, and only a few additional cases are lost due to measurement error or change of residence. As a result, 86 percent of the rural sample and 79 percent of the urban sample are used in the analysis that follows. We analyzed in detail lack of measurement in children’s height and found that, though missing data is not random, interactions of main parental characteristics with the relevant groups suggest there are no differential effects across the cohorts or locality groups used in the analysis.

Summary statistics (not shown) of the 2002 and 2005 height-for-age z-scores show that children 1 to 8 years old in Mexico are, on average, 0.56 standard deviations below the reference median in 2002 and 0.42 below the reference median in 2005 (standard deviations are 1.26 and 1.47, respectively). The percentage of stunted children, that is, children that are more than two standard deviations below the reference median, is 12 percent in both years.

As expected, the nutritional status of children in the rural areas is worse than that of children in the urban areas. Rural children are on average 0.75 and 0.59 below the reference median in 2002 and 2005 respectively, whereas the corresponding numbers for urban children are 0.42 and 0.29. With respect to stunting, the overall incidence of 12 percent is a combination of an incidence of stunting among rural children of 16 percent and an incidence of stunting in urban children of 10 percent.

---

8 Children that moved between 2002 and 2005 are eliminated from the 2005 sample because treatment status based on the place of residence does not correspond to actual treatment.
Two other sets of comparisons between the rural and urban areas are worth looking at in order to better interpret the results presented in the following sections. First, Oportunidades participation rates are much lower in urban areas. Two factors explain this: (i) the percentage of eligible households is smaller, and (ii) the take-up rate among eligible households is smaller. On average, rural communities have between 30 percent and 70 percent of the households enrolled in the program, while the urban communities have at most 17 percent (Table 2).

Second, due to the geographic targeting of the program, localities incorporated later in time tend to be relatively better-off. Using the Housing and Population and Count data we analyzed a broad set of socioeconomic variables, including dwelling characteristics (floor and roof materials, access to basic services such as running water and electricity), asset ownership, literacy rates, school attendance, and access to health insurance. Within rural areas, both the mean and the median (not shown) of each characteristic improve in communities that were incorporated later. In contrast, urban communities that received the program earlier look similar to those incorporated later. This is consistent with the fact that the geographic targeting of the program was done only in rural areas.9

4. Identification strategy

The identification strategy exploits the combination of three elements: (i) the evidence that nutritional interventions have only modest effects on child height after the first few years of life; (ii) the fact that Oportunidades was gradually expanded over the years; and (iii) the panel dimension of MxFLS. Based on the first two, program exposure is defined as a function of the age

---

9 Even though by looking at some socioeconomic indicators urban localities do not look, on average, better-off as they are incorporated later in time, the expansion was still not random. Therefore, we still think it is reasonable to keep working with the locality classification introduced earlier which uses year of entry to define locality groups.
of the child at the time Oportunidades was introduced to the locality of residence. The approach consists of identifying cohorts of children that were exposed to the program and cohorts of children that were not, and then performing an impact analysis at the community level. Variations of this empirical methodology have been successfully implemented in the nutrition and economics literatures (see, for example, Martorell and Habicht, 1986; Duflo, 2001; Frankenberg et al., 2005).

The biology of child growth defines a critical period over which nutritional interventions can significantly affect child height. While height is very responsive to nutritional inputs during the first years of life, it is largely determined after a given age. There is not absolute consensus in the literature regarding the exact age cut-off over which height is predetermined, but it does lie within between the two and four years of life. It has been shown that once children turn 4 years old, the influence of nutritional interventions on height is substantially reduced (Martorell and Habicht, 1986).

Based on this evidence, even though we could expect Oportunidades to potentially improve the nutritional status of all children, the improved nutritional status should only be manifested on height if children were young enough at the time Oportunidades arrived. Therefore, treated children are defined to be those who are younger than 4 years at the time the program arrived, and who are exposed to the program until they reach that age. Older children constitute our control group.10

In our main specification, we define exposure as a function of age, place of residence, and time of measurement. This means we follow an intent-to-treat approach, where every child in a given place will be treated as exposed if that child lives in a community that was covered by the program while

---

10 Throughout this section exposure is always defined in terms of height. Older cohorts clearly benefit from the program in other dimensions.
the child was younger than 4 years old. Yet, in the empirical analysis we complement this baseline specification with other specifications that restrict the sample of children to those more likely to have participated.

As mentioned before, the geographic expansion of Oportunidades was far from random. Therefore, we chose not to use as our control group children from places where the program had not arrived by the time of measurement. The reason is that it is very unlikely that controlling for observable locality characteristics will capture all of the locality-specific effects on height. Instead, we use the timing in the roll-out of the program to define relatively homogeneous groups of localities and perform an impact analysis within these locality groups. We do use, however, children in communities where the program does not operate to control for nation-wide time trends.

Using within-locality variation means that we need to identify impact effects from differences in the level of exposure between younger and older children. Transforming height into z-scores helps controlling for age-specific differences. However, we might still worry that the dynamics of height over the life course could contaminate the comparison between older and younger children at a given point in time. The third component of our identification strategy specifically addresses this concern. By measuring children at two points in time, 2002 and 2005, we are able to compare different cohorts of children but measured when they are at the same point over their lifecycle.

Taking into account our definition of exposure and the time of our survey, we divide our sample of children in three groups: a younger cohort, a middle cohort and an older cohort. The younger cohort \((y)\) consists of children between 1 and 3 years old in 2005 (born between 2001 and 2003). The middle cohort \((m)\) includes children 1 to 4 years old in 2002 and between 4 and 7 in 2005 (born between 1997 and 2000). The older cohort \((o)\) includes children between 5 and 7 years old.
in 2002, born between 1994 and 1996. We do not want to include children much older in the analysis, because once children reach puberty differences in height are hard to interpret. As a result, four cohort-time groups are used to identify treatment effects: the older and middle cohorts measured in 2002 and the middle and younger cohorts measured in 2005.\textsuperscript{11}

Next, localities are divided into different groups depending on the year they were incorporated into the program. We define four groups of rural localities and three groups of urban localities. In the rural areas the groups are the following: the first group consists of localities that were incorporated at the very beginning of the program, in 1997 or 1998; the second group consists of localities incorporated right after the first group, between the years 1999 and 2002; the third group includes the localities that received the program between 2003 and 2005; and finally the fourth group includes the localities that either received the program after 2005 or never did. Because the program only expanded to cover urban areas in 2001, only the last three groups of localities exist in the urban areas.\textsuperscript{12,13} Note that the selection of the groups is closely related to the two years in which height measures are taken.

At the time of measurement, the three cohorts of children will have experienced a different level of exposure to the program depending on the locality group they live in.\textsuperscript{14} To identify impact effects we are going to define three levels of exposure: Full Exposure refers to children exposed to the program since birth; Partial Exposure refers to children exposed to the program for some

\textsuperscript{11} In the estimation only a subgroup of the middle cohort is used in order to compare groups of children as similar as possible. More specifically, since the younger cohort in 2005 will be compared to the middle cohort in 2002, only children between 1 and 3 years old in 2002 are included in the analysis. Similarly, the middle cohort in 2005 will be compared to the older cohort in 2002, so only children between 5 and 7 years old in 2005 are included in the analysis.

\textsuperscript{12} Community group 2 is not exactly the same in the rural and urban areas, because in the urban areas this group includes three localities incorporated to Oportunidades in 1998 (See Figure 1). In terms of interpretation and exposition, however, they should be thought of as the same thing.

\textsuperscript{13} Results are robust to reclassifying the urban localities introduced to Oportunidades before 2001 as rural.

\textsuperscript{14} Implicit here is the assumption that the place of residence of our children at the time of the survey is the place of birth. We believe that this is a reasonable assumption, since migration is low for families with young children.
time before the age of 4 but not since birth; and Zero Exposure refers to children that were too old at the time Oportunidades arrived (4 or older). This information is presented in Table 3. Each cell represents the expected level of exposure to the program for a given cohort in a given locality-group at a given time.

Based on the given definition of cohorts and locality groups, we estimate the following regression equation:

$$\theta_{icvt} = \alpha_{ct}^4 + \alpha_{ct}^1 I_1 + \alpha_{ct}^2 I_2 + \alpha_{ct}^3 I_3 + \gamma'_c X_{ivt} + \epsilon_{ivt}, \quad (1)$$

where $\theta_{icvt}$ represents the height-for-age z-score of child $i$, in cohort $c$, in community group $v$, at time $t$. The superscript on each $\alpha$ coefficient represents the locality group. The specification allows for four different cohort-time intercepts ($\alpha_{ct}^4$ terms): an intercept for the middle and older cohorts at time zero (year 2002), and an intercept for the middle and younger cohorts at time one (year 2005). They estimate, for a given cohort-time group, the average height-for-age z-score of children in baseline or control communities. Each of these four intercepts interacted with a dummy that represents community-group 1 estimates, for the corresponding cohort-time group, the difference in average height-for-age of children in communities that received the program between 1997 and 1998 relative to children in baseline communities ($\alpha_{ct}^1$ terms). For instance, $\alpha_{c02}^1$ measures the difference, in 2002, between the average z-score of children in the older cohort in community-group 1 and the average z-score of children in the older cohort in baseline communities. $\alpha_{ct}^2$ and $\alpha_{ct}^3$ are interpreted in a similar manner. The set of covariates $X$ includes gender, age in months, presence of mother and father in household, height of mother and father,

---

15 In this section we refer to full or partial exposure broadly defined. In the empirical section we quantify potential exposure in months for each cohort in each locality group. This will help interpret the results.
16 The specification for the urban areas does not include the $\alpha_{ct}^1$ terms.
education of the mother, and state of residence. Note that the effect of each of these covariates is allowed to change across cohorts.\textsuperscript{17}

Table 4 shows estimated coefficients on a selective group of variables of equation (1). Panel A is a copy of Table 3, and presents the expected impact of the program as a function of the cohort of the child, the year height was measured and the type of locality of residence. Panel B shows the estimated results for rural areas, and Panel C presents the results for urban areas. Each cell presents the estimated coefficient corresponding to the $\alpha_{ct}^1$, $\alpha_{ct}^2$, and $\alpha_{ct}^3$ terms in equation (1). In Panel B, for example, the cell corresponding to the older cohort in community-group 1 in 2002 (-0.52) shows the estimated value of $\alpha_{o02}^1$, when equation (1) is restricted to the rural sample.

In the absence of selective program placement, the coefficients in equation (1) would already represent impact effects. But we know that is not the case. Looking at the rural areas, results in Table 4 are consistent with the fact that Oportunidades was first implemented in the poorest communities of Mexico, and reinforces descriptive analysis done with census data on a broader set of socioeconomic variables. As expected, rural communities incorporated to Oportunidades before 2005 are statistically different from communities incorporated later (or never incorporated) in terms of average children’s height: $\alpha_{o02}^2$ and $\alpha_{o02}^3$ are negative and significant at the 5 and 10 percent level, when they should be zero in the absence of selection in program placement. This analysis cannot provide direct evidence of placement selection for the first group of communities but the literature documents they are among the poorest communities in Mexico. As a result, we see that Oportunidades was introduced first to communities where children’s nutritional status was

\textsuperscript{17} We tried specifications that allow for the covariates to be cohort-time specific and they essentially produce the same results (except that some coefficients have larger standard errors).
worse off, and a simple comparison between children exposed and not exposed across communities will not provide unbiased estimates of program effects.

In contrast, in urban areas there is no evidence of selection in program placement at the locality level. The estimated coefficients of the older cohort at time zero are both close to zero and statistically insignificant. Once again, this is consistent with the geographic targeting policy described previously, as well as with the descriptive evidence performed with census data.

Therefore, to control for program placement, we use differences across $\alpha_{c\ell}$ terms for each locality group $\nu$. Because these terms are estimated relative to average z-score in control communities, these differences automatically eliminate any general time trend that makes our cohorts different from each other for reasons other than Oportunidades exposure. Finally, we control for the dynamics in height by comparing the middle cohort measured in 2005 with the older cohort measured in 2002 (5 to 7 years old at the time of measurement), or the younger cohort measured in 2005 with the middle cohort measured in 2002 (1 to 3 years old at the time of measurement).

**Measuring the impact for communities included between 1997 & 1998 (Group 1)**

Panel A in Table 4 shows that $(\alpha_{m05}^1 - \alpha_{o02}^1)$ gives an estimate of full relative to partial exposure. Because both coefficients are measured relative to children in baseline communities, the difference between them accounts for any locality-specific effect that can be expressed as an additive component common to both cohorts. Furthermore, we are comparing children 5 to 7 years old in both cases to account for any height-specific life-cycle effect. However, we are comparing different cohorts. This means that the estimate will also capture any time effect that might have existed between the time period the older cohort was in its critical growth years and the time period
the middle cohort was in its critical growth years. If these communities experienced economic
growth between those two periods, the middle cohort may have been exposed to a better
environment than the older cohort when they both were between 1 and 4 years old. Under those
circumstances, the difference between $\alpha_{m05}^1$ and $\alpha_{o02}^1$ would include both the additional exposure
to Oportunidades and the improvement over time that would have happened regardless of the
program.

If time trends between community groups 1 and 4 were parallel, the difference between these two
cohorts in control communities would take care of any relevant time effect. We can directly test
the importance of this effect assuming that the time effect is homogeneous across cohorts. We can
use the difference between the younger cohort in 2005 and the middle cohort in 2002, as these are
two groups of children that experienced the same level of exposure to the program by the time of
measurement and live in the same group of communities. As a result, $(\alpha_{y05}^1 - \alpha_{m02}^1)$ provides us
with an estimate of the time effect between the middle and younger cohorts in community group
1 relative to the same time trend in community group 4.

Assuming the relative difference across communities between the middle and younger cohorts is
a good indicator of the relative difference between the older and middle cohorts, the double
difference $(\alpha_{m05}^1 - \alpha_{o02}^1) - (\alpha_{y05}^1 - \alpha_{m02}^1)$ should give us an unbiased estimate of the program
effect under full exposure relative to partial exposure.\(^{18}\)

*Measuring the impact for communities included between 1999 & 2002 (Group 2)*

\(^{18}\) Note that the identification assumption regarding time effects is very specific. We need that the difference in the
environment between the time period the middle cohort was 1 to 4 years old and the time period the younger cohort
was 1 to 4 years old, be a good counterfactual for the difference between the time period the older cohort was 1 to 4
years old and the time period the middle cohort was 1 to 4 years old. The time effect between 2002 and 2005 is not
the relevant concept because the time trend between those years affects the height of the three cohorts differently.
Table 4 shows that the difference \((\alpha_{y05}^2 - \alpha_{m02}^2)\) provides with an estimate of full relative to partial exposure, and \((\alpha_{m05}^2 - \alpha_{o02}^2)\) gives an estimate of partial exposure. In this case, we cannot separately estimate differential time trends between community groups 2 and 4. In contrast to the other two groups of communities, we do not have two cohorts of children that experienced, at the time of measurement, the same level of exposure to the program. Consequently, the estimates will be capturing both program effects as well as any possible differential time effects that could exist between these two community groups.

**Measuring the impact for communities included between 2003 & 2005 (Group 3)**

Similarly, we can get an unbiased estimate of partial exposure for children in the latest community group. \((\alpha_{y05}^3 - \alpha_{m02}^3)\) estimates the combined effect of partial exposure to Oportunidades and differential time effects between locality groups 3 and 4. By comparing two groups of children that were not exposed to the program, \((\alpha_{m05}^3 - \alpha_{o02}^3)\) provides an estimate of differential time trends. Under the assumption that the time effect between the older and middle cohorts is a good counterfactual to the time effect between the middle and younger cohorts, the double difference \((\alpha_{y05}^3 - \alpha_{m02}^3) - (\alpha_{m05}^3 - \alpha_{o02}^3)\) gives an estimate of partial exposure to the program net of placement and time effects.

5. Robustness checks

**Intent-to-treat specification**

Our intent-to-treat approach has some clear advantages. First, we do not need to worry about self-selection bias due to program participation. In this paper we work with a definition of exposure that is exogenous to the household. Therefore, household-specific factors correlated with both
actual treatment status and the outcome of interest do not affect the estimates.\textsuperscript{19} Previous literature summarized above provides evidence of the importance to control for selective access and selective participation.

Second, the definition of treatment at the locality level makes the analysis robust to the existence of spillover effects. There is no evaluation assessing the existence of these effects on health outcomes, but there is evidence on other dimensions. Bobonis and Finan (2006) find that the program significantly affected school enrollment rates among non-beneficiary children and Angelucci and DeGiorgi (2009) find that Oportunidades increased food consumption among non-beneficiary households.

Nevertheless, the Intent-to-treat approach strategy has also some drawbacks. First of all, the treatment group that may include a very large proportion of children that did not benefit from the program. Since, the program targets poor households, children in better-off households are not eligible and not expected to benefit directly from this intervention.\textsuperscript{20}

To the extent that non-treated children cannot be made worse-off by the program, the estimated impact would provide a lower bound of the program effect on children’s height. However, the analysis may end up being uninformative if there is not enough power to identify positive impact effects. This would be the case if the share of untreated children in the cohort exposed to the program is sufficiently large. Given that Oportunidades was first introduced in more disadvantaged

\textsuperscript{19} Implicit here is the assumption that the time Oportunidades arrived in the locality of residence is exogenous to the household, which holds if Oportunidades did not induce migration from places without the program to places with the program. Note additionally that the fact that children’s height cannot be affected after they reach a certain age rules out the possibility that parents compensate untreated children. If that was the case, behavioral responses induced by program participation would need to be considered even if treatment status was exogenous to the households.

\textsuperscript{20} Every intent-to-treat effect estimator includes both treated and non-treated individuals. Note, however, that the share of non-treated individuals is expected to be higher in this case than in more traditional impact estimates. The reason is that this analysis not only classifies as treated those children that are eligible but decided not to participate or children that participate but have limited access, but also children not designed to participate.
places, the identification strategy as explained so far is expected to be weaker as we evaluate communities incorporated later in time. This effect is reinforced by the lower take-up rates among eligible households after 2001. To address this concern, we add to the baseline specification additional analyses that follow the same identification strategy, but restrict the sample in different ways. We continue to use variation in exposure between older and younger children, but only use the sample of children that live in households that, according to different criteria, are more likely to participate in the program.

**Time trends and community shocks**

Also important to the identification of treatment effects are the assumptions regarding time trends and shocks. The underlying assumption regarding time effects comprises two components: how time trends change across cohort groups, and how time trends differ across locality groups.

Regarding the first component, if different cohorts were exposed to different events or shocks, the estimated impact will not be able to disentangle the effect of the program from the effect of these events. This source of bias is partially controlled for by estimating differences across cohorts in treatment communities relative to control communities. However, if there are location-specific shocks that affect differently our exposed and non-exposed cohorts of children, these effects are not accounted for.

One concern is the fact that some localities (or households) were incorporated into Oportunidades responding to specific negative shocks, in most cases hurricanes. We use administrative information about those localities, and excluded them from the analysis as a robustness check.\(^\text{21}\)

---

\(^{21}\) Only 3 MxFLS localities were in this list, and all results are robust to the exclusion of these localities.
In addition, we use information on past shocks as reported in the MxFLS household and community questionnaires to check the robustness of our results to locality-specific events.

A related, though different, concern has to do with a possible bias arising from location-specific time trends. The regression equation controls for state-time fixed effects, and all effects are estimated relative to control communities. However, if there is permanent divergence or convergence in growth rates between our treatment community groups and the control community group, time effects are not completely accounted for. Also, one might worry that the relatively small number of control communities in our sample might affect our results. To address this concern, we also report results following an alternative identification strategy that does not use children in control communities. We estimate equation (1) for each locality group separately, which means that now impact effects are directly identified out of differences across the four cohort-time intercepts.\(^\text{22}\) The cost to following this alternative is that now we need somewhat stronger assumption to control for time trends within locality groups. For community groups 1 and 3, we need the difference between a given cohort and the younger (older) group to be a good counterfactual for the difference between that cohort and the older (younger) group. Furthermore, there is no way to control for time effects when we evaluate the second group of communities. One additional feature of this alternative is that now the analysis can be restricted to the sample of households who report being program participants.\(^\text{23}\)

In addition, Mexico experienced a financial crisis in 1994, known as the Tequila crises, which generated a widespread fall in income, consumption and wages. This is important for us because the group of children we labeled “older cohort” includes those children born between 1994 and

\(^{22}\) Due to sample size constraints, covariates are no longer cohort specific.

\(^{23}\) This is not possible under the main identification strategy, because by construction, there are no program participants in control communities.
1996. Two important points need to be made. First of all, as mentioned before, nation-wide business-cycle effects are captured by children in control communities. Additionally, we allow for time-specific geographic variation by having state-time fixed effects in our estimation. Secondly, the evidence suggests that the greatest effect was on households living in metropolitan areas, with highly educated household heads, and workers in financial services and construction. In contrast, the smallest effect was on less-educated, rural and agricultural workers (McKenzie, 2003). This suggests that our target population was relatively less affected by the crises. We should worry less about this issue when we estimate program effects for the rural areas, as well as when we restrict the urban sample to proxy the target population of this program. Still, we should keep this in mind as we read the results.

Finally, in 2003 the Mexican government introduced a program called Seguro Popular that provides with health insurance to households that do not have access to this service. The program expanded gradually over time, participation is voluntary, the only selection criteria is lack of health insurance, and the service is provided for free or at a very low cost. Fortunately for our purposes, participation rates in our localities are not very high (between 5 and 20 percent in rural areas and between 2 and 7 percent in urban areas in 2005). By 2010, participation rates are much higher. The fact that this program only offers an insurance product, when Oportunidades offers a wide set of benefits (including a sizable cash transfer) mitigates our concerns. However, results should be interpreted with the existence of this program in mind.
6. Results

Following the identification strategy presented in the previous section, we now present the estimated impact effects of the intervention on child nutrition. We focus first on the rural areas, and then on the urban areas.

Results for rural areas

Table 5 presenting the results for rural areas is divided in three panels. Panel A presents the estimates for the first group of rural communities incorporated between 1997-1998, the Panel B corresponds to community-group 2 that received the program between 1999-2002, and Panel C shows results for the last group of rural communities that entered between 2003-2005.

The first column in Table 5 presents the estimated results of our baseline specification, where we use all children in the rural areas. Looking at the poorest group of rural communities in Panel A, the evidence suggests that Oportunidades did have a positive effect on children’s height. The first row shows the difference in height between the middle cohort measured in 2005 and the older cohort measured in 2002, which would reflect the impact of being fully exposed to the program relative to being partially exposed to it, in the absence of differential time effects between this group of communities and control communities. If we look at the sample statistics, we see that the middle cohort was, by 2005, not only exposed to the intervention since birth, but it was also, on average, exposed for over two years more than the older cohort. Given the size of the difference in the degree of exposure, and the fact that these are very poor communities, it is not surprising that the estimated effect is positive and highly significant. Next, we look at the estimated time effect shown in row 2. The negative and statistically significant estimate indicates that the rate of improvement over time that took place in the first (and poorer) group of communities is slower
than that in control communities. Taking this into account, we get a higher estimate of “Full relative to Partial” exposure. According to this analysis, children fully exposed to the program are on average 0.88 standard deviations taller than children only partially exposed to it. This magnitude represents, for example, 3.7 cm for a four-year-old boy.

As explained in the previous section, an estimate of program exposure that is not confounded with differential time effects cannot be identified for the second group of rural communities (Panel B). Keeping this in mind, we don’t see evidence of a positive impact effect on these children. Likewise, we don’t see any positive effect on the nutritional status of children living in the group of rural communities incorporated at last (Panel C). The estimated effect of partial exposure, with or without accounting for differential time trends, is negative and statistically insignificant.

As mentioned before, our baseline specification includes all children in the analysis, even though only a subsample of them is expected to benefit from the intervention. For this reason, we restrict the sample in different ways and show the results in columns (2) to (4). Column (2) presents the estimates when we exclude from the analysis better-off children measured by their mother’s education - those whose mothers have completed high-school education or more (12 or more years of education). In column (3) we further restrict the sample to those children whose mothers have at most primary education. Finally, in column (4) we implement the Oportunidades eligibility criteria and only keep the sample of children who live in program-eligible households. That is, we only keep children living in households whose poverty score is above the official program cut-off.

Looking at the poorest group of communities (Panel A) the story remains unchanged, except that now impact estimates are more imprecisely estimated when using mother’s education to select the sample. Column (2) shows the exact same point estimate we get in our baseline specification, but
the standard error is larger and the effect is not significant. If we look at column (3), the point estimate is smaller and insignificant. However, if for efficiency reasons we restricted the -0.05 time effect to be zero, the results would suggest a positive and significant impact on these children. When we focus on program-eligible children, the effects are bigger and highly significant.

Next, we analyze the second group of communities (Panel B). Since this group is relatively better-off than the first one, we might expect to see some differences once we focus on more disadvantaged children. Results in columns (2) to (4) confirm this intuition. The table shows that the estimated effect of partial exposure is now larger and statistically significant. This suggests that, unless differential time effects for this community group were positive and sufficiently big, Oportunidades seems to have had a positive effect on these children’s height.

In contrast, the bottom panel suggests that conclusions do not change when we analyze communities incorporated after 2002 (Panel C). Regardless of the sample stratification used, there is no evidence of a positive impact on these children. Even though impact effects are not statistically significant, point estimates for the last group of communities are negative and quite large.

We search for evidence on negative shocks that might be behind the estimates, particularly for the third group of localities. As mentioned in the previous section, we obtained from Oportunidades records a list of localities in which some households were enrolled in the program through a special process after having been affected by a natural disaster. This would bias our estimates downwards, as there would be perfect correlation between program entry (positive shock) and the negative shock. It turns out that only 3 MxFLS localities are in this list, and all results are insensitive to including them or not.
Additionally, we looked for evidence on negative shocks within our survey. Both the community and the household questionnaires have a section that asks about specific shocks experienced by the community (household) during the 3 years (5 years) previous to the interview. Using this information we found that in a few localities the percentage of households that reported having lost their house, business or total crop production was as high as 30 to 40 percent. To confirm that our results are not driven by any of these particular places we run the regressions without these localities and all results remain unchanged. We also run the models using the entire sample but adding a dummy variable that takes the value one if the community questionnaire indicates that the locality suffered any shock in the previous three years, and that dummy interacted with each locality group. Once again, none of the results change either in the rural or urban areas.

Finally, to make sure results are not driven by the relatively small sample of children in control communities, by idiosyncratic effects in our control communities, or by differential time effects across community groups, we turn next to estimates that analyze each community group separately. In column (5) we use all the children in our sample, and then in columns (6) to (8) we restrict the sample in different ways as before. In addition, in column (9) we show the estimates if we use only children in households who self-report being in the program at the time of the survey. We can only do this for the first two groups of communities, because by construction, there are no children in participating households in community-group 3 in 2002. For the same reason we do

\[24\] At the locality level the enumerated shocks include flood, earthquake, landslide, fire, hurricane, drought, plague, frost and hailstorm. At the household level, the questionnaire asks about the death or illness of household members, unemployment, and crop, production animals or property losses. This information is not complete, as we still lack information on self-reported shocks at the time the older cohort was going through its critical years (0-4 years old). Still, information reported in 2002 is relevant for the middle cohort, and information reported in 2005 is relevant for the younger cohort. This is important considering the only negative estimated effect corresponds to the last group of rural communities, where the main impact is identified from differences between the younger and middle cohorts.
not use this sample stratification when using the baseline specification – there are no children in participant households in control communities.

The message is remarkably similar across columns. For the poorest group of communities, point estimates are around 0.17 and 0.28. Even though net estimates of exposure are not statistically significant, except for column 6, they would if we restricted the small point estimates of time effects to be zero. For the second group of communities, the estimated impact of partial exposure remains positive, and for the most part significant. In addition, now the estimated impact of full relative to partial exposure is larger and significant. Again, we still do not see evidence of a positive impact on communities incorporated after 2002.

**Results for urban areas**

The situation in urban areas is quite different than that in rural areas. First, heterogeneity in socio-economic status is much higher, and the percentage of program-eligible households much lower. Additionally, the evidence suggests that participation rates among eligible households are much lower. Based on this evidence, we do not expect our baseline specification, which uses all the sample of children, to be very informative. Column (1) in Table 6 confirms this intuition. Contrary to what one might expect, it looks like children exposed to the program are actually made worse-off. While the estimated effect of partial exposure for children in the second group of communities is positive and significant, the other two impact effects are negative and significant. Since we do not expect the program to affect child nutrition in a negative way, we suspect this estimates are either driven by better-off households, or program exposure is correlated with other negative shocks. We analyze these hypotheses next.
Columns (2), (3) and (4) show impact effects when we eliminate from the analysis better-off households. Just as in the rural areas, the first two columns restrict the sample based on the education level of the mother, and column (4) imposes the Oportunidades eligibility criteria to select beneficiary households. As expected, restricting the sample in any of these ways substantially reduces the urban sample size. Approximately 80 percent of the children have mothers with less than 12 years of education and this number falls to 35 percent if we only consider those with mothers that have at most primary education. When looking at program-eligible households, only 16 percent of the original sample remains. None of the results shows evidence of a positive impact on child height in either locality group. It is remarkable how this result holds even after focusing on the most disadvantaged groups.

Regarding looking for evidence on negative shocks, we performed the same analyses mentioned for the rural areas. Neither administrative records nor self-reported negative shocks can explain the negative coefficients shown in Table 6.

We turn next to estimates that analyze each community group separately. These results are shown in columns (5) through (9), where the first column uses the complete sample, and the rest of the columns restrict the sample in the ways previously described. When using this alternative identification strategy, there is some evidence of a positive impact of the program on child height for children in communities incorporated before 2002. Table 6 shows that the estimated effect of partial exposure is positive and significant. Recall though, that when we use this identification strategy in community-group 2 no time effect is controlled for. The estimated coefficient is capturing both, the exposure to the program as well as any improvement that might differentiate the middle and older cohorts. The situation is different when we estimate the effects using all locality groups, as it is only differential time trends that are not taken into account. Finally, we still
do not see any evidence of a positive impact when we evaluate the last community group. Partial exposure, net of time trends, remains negative and insignificant.

7. Discussion

In order to put our results in context, we compare the estimates to those found in previous studies. (See Appendix 1 for a more detailed explanation of the previous literature.) As mentioned before, the Oportunidades rural evaluation sample covers localities incorporated in 1998. Therefore, we can compare our results for the first group of rural communities to the existing evidence in the rural areas. Experimental evidence suggests that children 12 to 36 months at baseline were about 1 cm taller after one year in the program. Analyses that evaluate one additional year of exposure after both treatment and control groups were in the program, only find positive effects on children living in more disadvantaged households (1 to 1.5 cm). We choose to convert our results for a 2 year-old boy, as it lies in between the 12 to 36 month range of previous estimates. Our baseline estimated effect of 0.88 standard deviations translates into 2.68 cm. Our second specification shows estimates in the range of 0.18 to 0.28 standard deviations. A middle point, 0.23 sd., translates into 0.7 cm. Given that these estimates represent an average additional exposure of two years, under the assumption of linear effects we get an estimated impact between 0.35 cm and 1.34 cm for an additional year. As we see, this range is consistent with previous evidence.

Even though our results cover a wide range of values, our preferred specification for this locality group is the one that uses all children and all locality groups in the estimation. By using all locality groups we can control for nation-wide time trends, and by including all children we are maximizing the sample of children in control communities without going too far away from the target population in community-group 1, as participation rates are very high and variation in socio-
economic conditions is low in this locality group. If we take this as our main result, our impact effect is somewhat bigger than previous evidence. Among other things, this could be due to the fact that our cohort of children fully exposed to the program was, not only exposed since birth, but also exposed for 4 years. Therefore, this group of children experienced the maximum level of exposure one could get from this intervention, including pre-birth exposure for some of them. Additionally, it could be that the documented problems in the initial implementation of the nutritional component of the program were fixed over time, or that there is a learning process happening at the household or community level. We are measuring impact effects seven years after the program started.

The only existing evidence from the urban areas analyzes the effect after two years of exposure on children incorporated in 2002 (Leroy et al., 2008). The study only finds a positive effect on height for children who were younger than 6 months at baseline (0.41 z-score). When the authors divide the sample by socioeconomic status, they find a positive effect (0.27 z-score) on children in the poorest tercile. The closest estimate we have to compare this result with is our “Full relative to Partial exposure” estimate in community-group 2. In that case, we are approximately estimating the same length of exposure – this estimate represents an average of two additional years of exposure. Additionally, since children exposed for longer are in the program since birth, they can be closer compared with the subsample of children for which previous estimates found positive effects (younger than 6 months). Finally, even though we are comparing children fully exposed to children partially exposed (as opposed to children not exposed), our middle cohort of children was

25 This exposure is not on average. Every child in this cohort was exposed for four years, as virtually all are born after the program is in place, and they are all measured in 2005 when they are over 5 years old.
26 Another difference could come from sample composition effects. The Oportunidades data sampled households in 1997, while our survey represents the population of these localities in 2002.
on average, only exposed for about 4 months. Table 6 shows that this estimate is never significant, even when we focused on more disadvantaged households – most of the time is negative, and when positive the point estimate is quite small. Our estimates for the last group of communities also uses children exposed since birth, although for a shorter period of time – on average, about 15 months. These estimates do not change as we focus on more disadvantaged groups.

Probably, the biggest limitation of this work is that we cannot focus too narrowly on specific subgroups due to sample size constraints. Actually, our estimates are in general poorly estimated, especially in the urban areas, as our point estimates are not very stable and standard errors are quite high. Therefore, it could be argued that there might be positive effects on beneficiary households that this analysis is not capturing. Even though that is entirely possible, it should be noted that the only evidence we have in the urban areas implements standard non-experimental methods, is not representative of the beneficiary population incorporated in 2002, it has to deal with very high attrition rates, and only finds effects on very specific subgroups of the beneficiary population. We want to emphasize that the reason this estimation strategy is arguably weaker in the urban areas is due to the very low participation rates, and this is precisely the biggest concern with current estimates of the program impact in urban areas, as self-selection into treatment potentially constitutes an important source of bias. All in all, we conclude that although it is possible the program has a positive impact on child height, up to this day there is no conclusive evidence that supports this.

An alternative claim that could be made is that we are only looking at one marker of nutrition - height. Then, it could still be the case that the program has positive effects on other nutritional indicators. While this is entirely possible, previous evidence points at an ever weaker effect of the program on other outcomes. Leroy et al. (2009) summarize the evidence on nutrition for 5 CCT
programs in Latin America, including Oportunidades. They show how anthropometrics have been found to be the place where CCTs have the highest impact, and within anthropometrics, height has been found to be more responsive than weight. Evidence on micronutrient status summarized in that paper shows modest impact of Oportunidades on hemoglobin and anemia prevalence, and no impact on other micronutrients such as vitamin A, iron or zinc.

Assuming we believe the program does not have an effect on height on children living in rural communities incorporated after 2002 and in urban places, finding the reasons why this is the case is not straightforward. Even though these communities are relatively better-off, we are not looking at an outcome for which there is no room for improvement. Therefore, we should expect the intervention to improve the nutritional status of participating children.

As we mentioned in Section 2, implementation problems have been documented both in the rural and urban areas. There is evidence that nutritional supplements were not consumed as indicated, both in terms of frequency as well as preparation. Supplements were shared with other household members, and the meetings were not effective at changing households’ habits. Additionally, it was discovered that the iron in the supplement was not easily absorbed. For all these reasons, a series of changes were introduced later: it was better stressed the need to target supplements to intended children, the formula for the supplement was improved (for better iron absorption and other micronutrients were added), and the way community meetings were held changed (to stimulate more participation as opposed to being a lecture) (Neufeld et al., 2005). These changes were incorporated after September 2005, so they are outside our sample period. Therefore, poor implementation could partially explain the lack of impact. However, to explain the results these problems need to be a bigger issue in communities incorporated later. Also, given the size of the monetary transfer, one would expect some effect even if the nutritional component is not perfectly
implemented. The fact that the transfer represents a smaller percentage of expenditures for urban households might also help explain the results.

8. Conclusion

In this paper we are able to exploit unique panel data to inform about the impact of Oportunidades on child nutrition. We contribute to the literature by measuring the effects using a representative survey that allows to assess the impacts in areas that were incorporated after the first implementation phase, which was focused on very poor rural communities. This is the only evaluation that relies on survey data outside the Oportunidades Evaluation samples. This analysis aims to inform an emerging literature about the potential pitfalls of scaling up programs which may be linked to changes in the characteristics of beneficiaries or differences in the context (Al-Ubaydli et al. 2019).

Our analysis focused on child nutrition. Nutritional status could be measured in a number of ways. In this project, we chose height as our marker of interest for a number of reasons. On the one hand, there is a well-establish positive association between adult height and earnings. Controlling for a number of individual characteristics, taller individuals earn, on average, higher earnings, and this relationship has been found to hold across a wide variety of contexts. On the other hand, we know from the nutritional literature that height at age two is a very strong predictor of adult height. Also, it has been shown that height is sensitive to nutritional inputs during the first years of life but not later in life. This feature is key for our identification strategy. Finally, height constitutes a long-term marker of nutritional status, as opposed to other indicators that measure short-term nutritional status. In summary, we have a marker that is responsive to nutritional interventions early in life (better nourished children are, on average, taller), a marker that measures long-term nutritional
status (as opposed to short-term nutritional status), and a marker that has a well-defined welfare meaning (given by its link to adult earnings, among other outcomes).

Just as with any non-experimental study, the main challenge in the paper is to be able to isolate impact effects in the absence of experimental variation. To overcome this difficulty, we propose an identification strategy that combines insights from the biology of child growth, the timing of the roll-out of Oportunidades and the panel dimension of MxFLS. The strategy basically consists of identifying cohorts of children that were exposed to the program and cohorts of children that were not and then performing an impact analysis at the community level.

The analysis presented suggests that Oportunidades seems to be having a positive effect on child height in children living in the poorest communities of rural Mexico. However, we do not see evidence of a positive impact on rural communities incorporated after 2002. Similarly, there is no strong evidence of a positive impact on urban children. Estimates are always negative in communities incorporated after 2002, and they are only positive (and significant) in communities incorporated earlier for the specification that does not control in any way for time effects.

Overall, the findings of this paper show that when program placement is targeted or implemented in phases, initial impact results may not necessarily inform about the programs’ performance in other regions or populations. Moreover, we highlight the challenges in evaluating the impact of antipoverty interventions, even with an outcome such as child height which allows for a more rigorous assessment of impacts. Improving the quality of data collection to systematically assess long-run impacts and program sustainability is also central to support the expansion of highly costly programs such as Oportunidades.
References


9. Tables and Figures

Figure 1. Expansion of Oportunidades over time at the locality level

![Figure 1. Expansion of Oportunidades over time at the locality level](image-url)
Table 1. Final sample used in the analysis, by area.

<table>
<thead>
<tr>
<th>Children 1 to 3 and 5 to 7 years old</th>
<th>2002</th>
<th>2005</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>#</td>
<td>%</td>
<td>#</td>
</tr>
<tr>
<td>Total number of observations</td>
<td>4542</td>
<td></td>
<td>3807</td>
</tr>
<tr>
<td>Total number of observations in rural areas</td>
<td>1931</td>
<td></td>
<td>1569</td>
</tr>
<tr>
<td>Observations lost due to:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>missing height</td>
<td>307</td>
<td>172</td>
<td></td>
</tr>
<tr>
<td>+ missing z-score</td>
<td>0</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>+ moved</td>
<td>19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Final rural sample</td>
<td>1624</td>
<td>84%</td>
<td>1375</td>
</tr>
<tr>
<td>Total number of observations in urban areas</td>
<td>2611</td>
<td></td>
<td>2238</td>
</tr>
<tr>
<td>Observations lost due to:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>missing height</td>
<td>496</td>
<td>492</td>
<td></td>
</tr>
<tr>
<td>+ missing z-score</td>
<td>0</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>+ moved</td>
<td>22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Final urban sample</td>
<td>2115</td>
<td>81%</td>
<td>1719</td>
</tr>
</tbody>
</table>

If children moved between 2002 and 2005 height is set to missing in 2005 (the observations are treated as if they were individuals not found in 2005)
Source: MxFLS 2002 and 2005

Table 2. Oportunidades participation rates by locality groups

<table>
<thead>
<tr>
<th>COMMUNITY GROUP 1</th>
<th>RURAL AREAS</th>
<th>URBAN AREAS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incorporated between 1997 and 1998</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>p25 45</td>
<td>p50 57</td>
</tr>
<tr>
<td>2005</td>
<td>45</td>
<td>58</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>COMMUNITY GROUP 2</th>
<th>RURAL AREAS</th>
<th>URBAN AREAS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incorporated between 1999 and 2002</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2005</td>
<td>40</td>
<td>50</td>
</tr>
<tr>
<td>COMMUNITY GROUP 3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Incorporated between 2003 and 2005

<table>
<thead>
<tr>
<th></th>
<th>2000</th>
<th>2005</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>7</td>
</tr>
</tbody>
</table>

Source: Oportunidades administrative data

Table 3. Expected impact of Oportunidades on children’s height as a function of their cohort, locality of residence and year of measurement, by rural-urban sector

<table>
<thead>
<tr>
<th>Height measured in:</th>
<th>2002</th>
<th>2005</th>
<th>2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>Locality group:</td>
<td>Group 1</td>
<td>Group 2</td>
<td>Group 3</td>
</tr>
<tr>
<td>Panel A: Rural</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Old Cohort (o)</td>
<td>Partial</td>
<td>Zero</td>
<td>Zero</td>
</tr>
<tr>
<td>Middle Cohort (m)</td>
<td>Full</td>
<td>Partial</td>
<td>Zero</td>
</tr>
<tr>
<td>Young Cohort (y)</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>Panel B: Urban</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Old Cohort (o)</td>
<td>.</td>
<td>Zero</td>
<td>Zero</td>
</tr>
<tr>
<td>Middle Cohort (m)</td>
<td>.</td>
<td>Partial</td>
<td>Zero</td>
</tr>
<tr>
<td>Young Cohort (y)</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>
**Table 4. Estimated impact of Oportunidades on children's height by type of locality and rural-urban sector**

<table>
<thead>
<tr>
<th>Locality group:</th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Height measured in:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panel A: Expected impact</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Old Cohort (o)</td>
<td>Partial</td>
<td>Zero</td>
<td>Zero</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>Middle Cohort (m)</td>
<td>Full</td>
<td>Partial</td>
<td>Zero</td>
<td>Full</td>
<td>Partial</td>
<td>Zero</td>
</tr>
<tr>
<td>Young Cohort (y)</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>Full</td>
<td>Full</td>
<td>Partial</td>
</tr>
<tr>
<td>Panel B: Rural areas</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Old Cohort (o)</td>
<td>-0.52</td>
<td>-0.46</td>
<td>-0.28</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>Middle Cohort (m)</td>
<td>-0.31</td>
<td>-0.41</td>
<td>-0.27</td>
<td>-0.19</td>
<td>-0.24</td>
<td>-0.04</td>
</tr>
<tr>
<td>Young Cohort (y)</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>-0.860</td>
<td>-0.470</td>
<td>-0.690</td>
</tr>
<tr>
<td>Panel C: Urban areas</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Old Cohort (o)</td>
<td>.</td>
<td>-0.08</td>
<td>0.03</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>Middle Cohort (m)</td>
<td>.</td>
<td>0.35</td>
<td>0.29</td>
<td>.</td>
<td>0.31</td>
<td>0.21</td>
</tr>
<tr>
<td>Young Cohort (y)</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>-0.23</td>
<td>-0.11</td>
<td>.</td>
</tr>
</tbody>
</table>

Robust standard errors in squared brackets clustered at the locality level. Reference category: communities that did not have Oportunidades by 2005. Regressions control for: gender of child, age in months, presence of mother and father in the household, mother's and father's height, mother's education, state of residence.

**Table 5. Estimated impact of Oportunidades on child height. Rural Areas.**

<table>
<thead>
<tr>
<th>Sample of children</th>
<th>Baseline Specification. All locality groups</th>
<th>Each locality group separately</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ENTIRE</td>
<td>MOTHER'S EDU&lt;12 yrs</td>
</tr>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td><strong>Panel A: Locality group 1</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FULL relative to PARTIAL exposure (m05-o02)*</td>
<td>0.34**</td>
<td>0.41**</td>
</tr>
<tr>
<td></td>
<td>[0.18]</td>
<td>[0.19]</td>
</tr>
<tr>
<td>TIME effects (y05-m02)</td>
<td>-0.55*</td>
<td>-0.46</td>
</tr>
<tr>
<td></td>
<td>[0.32]</td>
<td>[0.49]</td>
</tr>
<tr>
<td>FULL relative to PARTIAL exposure</td>
<td>0.88**</td>
<td>0.88</td>
</tr>
<tr>
<td></td>
<td>[0.41]</td>
<td>[0.59]</td>
</tr>
<tr>
<td><strong>Panel B: Locality group 2</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FULL relative to PARTIAL exposure (y05-m02)*</td>
<td>-0.06</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>[0.37]</td>
<td>[0.47]</td>
</tr>
<tr>
<td>PARTIAL exposure (m05-o02)*</td>
<td>0.22</td>
<td>0.47**</td>
</tr>
<tr>
<td></td>
<td>[0.19]</td>
<td>[0.22]</td>
</tr>
<tr>
<td><strong>Panel C: Locality group 3</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TIME effects (m05-o02)*</td>
<td>0.24</td>
<td>0.44*</td>
</tr>
<tr>
<td></td>
<td>[0.20]</td>
<td>[0.24]</td>
</tr>
<tr>
<td>PARTIAL exposure (y05-m02)*</td>
<td>-0.43</td>
<td>-0.27</td>
</tr>
<tr>
<td></td>
<td>[0.32]</td>
<td>[0.50]</td>
</tr>
<tr>
<td>PARTIAL exposure</td>
<td>-0.67</td>
<td>-0.71</td>
</tr>
<tr>
<td>------------------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td></td>
<td>[0.44]</td>
<td>[0.64]</td>
</tr>
</tbody>
</table>
| Total Observations | 3003 | 2813  | 1852  | 2077  | 1454  | 1407  | 1029  | 1137 | 955  
| Locality group 1  |       |       |       |       |       |       |       |      |
| Locality group 2  |       |       |       |       |       |       |       |      |
| Locality group 3  |       |       |       |       |       |       |       |      |

Notes: Standard errors in brackets clustered at the locality level. * Significant at 10%; ** Significant at 5%; *** Significant at 1%; + Includes time effects


<table>
<thead>
<tr>
<th>Sample of children</th>
<th>Baseline Specification. All locality groups</th>
<th>Each locality group separately</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ENTIRE SAMPLE</td>
<td>MOTHER'S EDU&lt; 12 yrs</td>
</tr>
<tr>
<td>Panel A: Locality group 1</td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Panel B: Locality group 2</td>
<td>FULL relative to PARTIAL exposure (y05-m02)⁺</td>
<td>-0.58**</td>
</tr>
<tr>
<td></td>
<td>[0.24]</td>
<td>[0.22]</td>
</tr>
<tr>
<td>PARTIAL exposure (m05-o02)⁺</td>
<td>0.39**</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>[0.19]</td>
<td>[0.20]</td>
</tr>
<tr>
<td>Panel C: Locality group 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.18</td>
<td>-0.01</td>
</tr>
<tr>
<td>------------------</td>
<td>---------</td>
<td>--------</td>
</tr>
<tr>
<td></td>
<td>[0.22]</td>
<td>[0.22]</td>
</tr>
<tr>
<td>TIME effects</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(m05-o02)$^+$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PARTIAL</td>
<td>-0.40</td>
<td>-0.18</td>
</tr>
<tr>
<td>exposure</td>
<td>[0.27]</td>
<td>[0.25]</td>
</tr>
<tr>
<td>(y05-m02)$^+$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PARTIAL</td>
<td>-0.58*</td>
<td>-0.17</td>
</tr>
<tr>
<td>exposure</td>
<td>[0.33]</td>
<td>[0.30]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>3828</th>
<th>3160</th>
<th>1331</th>
<th>622</th>
<th>2585</th>
<th>2147</th>
<th>900</th>
<th>467</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Observations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Community group 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Community group 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: Standard errors in brackets clustered at the locality level. * Significant at 10%; ** Significant at 5%; *** Significant at 1%; $^+$ Includes time effects
N/A Does not apply to urban areas
Appendix 1. Existing evidence on the program’s impact

Evidence for rural areas

At the time the program started, an experimental design was set up to conduct a rigorous impact evaluation. Among a sub-group of the rural communities deemed eligible to receive the program in 1998, a group was randomly selected to receive the intervention right away while the other group was assigned to start 18 months later. By the year 2000 both groups were already under treatment. A comprehensive household survey was collected in these communities at baseline in 1997, followed by a series of survey rounds collected every six months until the year 2000. This survey is known as the “Oportunidades Rural Evaluation Sample”. This randomized controlled experiment was the foundation of an extensive body of work that evaluates the impact of Oportunidades on a myriad of outcomes, including consumption, wealth, labor market, education, and health related outcomes.

Three papers use the experimental design to evaluate the short term effect of the program on child height after one year of exposure. Gertler (2004) finds that children 12 to 36 months old in 1999 in treatment villages are on average 1 cm taller than children in control villages. Behrman and Hoddinott (2005) find similar effects but only when the authors control for the selective access to nutritional supplements driven by the shortage of supplements during the first years of the program. Rivera et al. (2004) find an impact of same magnitude when comparing children with

---

27 Other health outcomes evaluated in the literature include obesity, anemia, weight-for-height, BMI-for-age, birth weight, probability of illness.
28 The authors report that during the first years of the program, between 52 and 63 percent of children aged 4 to 48 months had access to the supplements. Adato et al. (2000) report that the distribution and intake of nutritional supplements seem to have been the most serious operational problem of the health component of Oportunidades. In response, health personnel exercised some discretion in the distribution of supplements by targeting those children
two relative to one year of exposure, but only among children who are 6 months or younger at baseline and live in the poorest households.

The main limitation of the Rural Evaluation Sample is that it is only representative of the communities incorporated to Oportunidades in 1998 and that only allowed for the estimation of very short term effects. To partially address the second issue, two additional rounds of data were collected in 2003 and 2007. The 2003 round incorporates a new non-experimental control group which was selected based on matching locality-level characteristics. This allows for the estimation of medium-term effects, but estimates need to rely on non-experimental identification strategies. This brings three main challenges: (i) the fact that the new control group resides in different geographic areas and therefore location-specific unobserved factors may affect comparisons; (ii) the fact that pre-intervention information on the new control group is elicited based on recall questions about their situation back in 1997 (introducing recall bias in diff-in-diff estimates); and (iii) the likely selective attrition in the original sample which is largely explained by changes of residence or migration (Teruel and Rubalcava, 2007).

Based on matching estimates using the 2003 data, Neufeld et al. (2004a) find that children 24 to 71 months old in 2003 in the original evaluation sample grew on average 0.67 cm more than children in the new control communities. Taking advantage of the original experimental design, that presented the most severe malnutrition symptoms (who could be in control areas). Similar intake problems are reported by Neufeld et al. (2004b) in the urban areas.

29 Three additional shortcomings with the experimental evidence are worth mentioning. First, the randomization was done at the village level, while the analysis of impact is at the individual level. Behrman and Todd (1999) find statistically significant differences in some pre-program characteristics when evaluated at the household or individual level. Second, attrition between rounds is not controlled for, and the work by Teruel and Rubalcava (2007) find that treatment households are more likely to leave the sample by the year 2000 relative to control households. Finally, indicators of nutritional status were not collected as part of the general evaluation survey, but collected at different times and by different teams under the supervision of the National Institute of Public Health. This implementation protocol introduced complications when trying to link the two data sources and the timing was such that baseline nutrition indicators in some households were measured after they had received program transfers.
the authors also estimate the impact of differential exposure between children in the original control and treatment communities (children 24 to 71 months in 2003) and find no significant effect. Similarly, Fernald et al. (2009) do the same on children born between March 1997 and October 1998 and measured in 2007. They only find a significant effect of about 1.5 cm on height in children whose mothers have no formal education.

**Evidence for urban areas**

When Oportunidades expanded to urban areas an “Urban Evaluation Sample” was collected. However, in contrast with the rural sample, the urban evaluation design was not experimental. In 2001, a sample of poor blocks from the areas assigned to receive the program in 2002 was selected. The control group was chosen based on a matching process from localities planned to be incorporated to the program in 2004. The data consists of a baseline survey fielded in 2002 and two follow-ups in 2003 and 2004. To analyze nutritional outcomes, however, a different control group was selected. In order to save on costs, control children were identified as those children in eligible households in treatment communities but who had not enrolled in the program. Adding to the self-selection concerns, these data suffer from an attrition rate of 40 to 45 percent.

Based on this two-year panel, and relying on a difference-in-difference propensity score matching estimator, Leroy et al. (2008) evaluate the impact of the program on children who were younger than 24 months at baseline. They find that after two years of exposure, there is no significant impact on growth in children 6 to 24 months at baseline, but there is a positive impact on children younger than 6 months old. The height-for-age z-score of treated children in this younger group is 0.41 higher than that of control children.