American Economic Journal: Economic Policy 2013, 5(1): 146–178
http://dx.doi.org/10.1257/pol.5.1.146

The Demand for Food of Poor Urban Mexican Households: Understanding Policy Impacts Using Structural Models[†]

By Manuela Angelucci and Orazio Attanasio*

We use Oportunidades, a conditional cash transfer to women, to show that standard demand models do not represent the sample's behavior: Oportunidades increases eligible households' food budget shares, despite food being a necessity; demand for food and high-protein food changes over time only in treatment areas; the treatment effects on food and high-protein food consumption are larger than the prediction from the Engel curves at baseline; and the curves do not change in eligible households with high baseline bargaining power for the transfer recipient. Thus, handing transfers to women is a likely determinant of the observed nutritional changes. (*JEL* D12, H23, J16, O12)

In this paper, we estimate a model of consumer demand using data on poor urban Mexican households. We then use such a model to predict the impact on the composition of consumption of a policy intervention that, in some areas, transfers cash to the women of eligible households. Finally, we compare the predictions of our structural model on expenditure shares with the impacts of the program we estimate by using the quasi-experimental variation induced by the program, and state-of-the-art econometric techniques which combine propensity score matching with difference in differences. Our exercise, therefore, constitutes an important test of the validity of the demand model. As we report some important rejections, we also suggest some plausible explanations for these rejections.

Consumer demand models—an important part of applied welfare economic analysis—are standard tools used to evaluate the welfare effects of tax policy reforms as well as changes in other circumstances faced by consumers. Such models, which can be used to analyze changes in direct and indirect taxation, in relative prices (induced by market forces or by subsidies), in benefit systems, or in transfer schemes, have been widely estimated using data from both developing and developed countries. The literature is too vast to be cited here.

*Angelucci: University of Michigan, Department of Economics, 611 Tappan Street, Ann Arbor, MI 48109 and IZA (e-mail: mangeluc@umich.edu); Attanasio: University College London, Economics Department, 30 Gordon Street, London, WC1E6BT and NBER, BREAD, and CEPR. (e-mail: o.attanasio@ucl.ac.uk). We would like to thank Erich Battistin, Jere Behrman, Jonah Gelbach, Kei Hirano, Costas Meghir, Barbara Sianesi, Petra Todd, Marcos Vera-Hernández, and Iliana Yaschine for very useful discussions of the issues covered by this paper, and an anonymous referee for many useful comments. Several staff at *Oportunidades* were very helpful with questions about the details of the data. Attanasio's research was partly funded by a European Research Council Advanced Grant No. AdG -249612-IHKDC.

[†] To comment on this article in the online discussion forum, or to view additional materials, visit the article page at http://dx.doi.org/10.1257/pol.5.1.146.

To be used for welfare analysis, demand systems have to be consistent with consumer theory. Indeed, one cornerstone of this literature is the derivation of estimable demand systems that are theory-consistent. Deaton and Muellbauer (1980), for instance, propose the Almost Ideal Demand System (AIDS), which Banks, Blundell, and Lewbel (1997) generalize to Quadratic AIDS (QUAIDS). These demand systems are typically derived from a unitary household model, in which the household is considered a decision unit and possible differences between household members are ignored. An important implication of this approach is that budget shares depend only on variables that affect preferences (such as family composition) or budget constraints (such as total expenditure or prices), and not on variables that affect exclusively the distribution of resources within the family (such as the share of resources controlled by a household member).

A large body of evidence points to the rejection of the unitary model in favor of models in which agents have different preferences.[1] Most of these papers show that household demand is significantly affected by variables that should not determine demand under the unitary model. Browning and Chiappori (1998), for example, show that budget shares depend on the wife's share of household income, conditional on total consumption.

This literature, however, has two shortcomings. First, some of these "rejections" of the unitary model can actually be rationalized as changes in the preferences of the unitary household. This is especially the case when one tests the unitary household model using policy interventions that change both relative incomes and preferences in the household. Any shift of the demand curve could be attributed to a change in the decision-making power of the household member who controls a larger fraction of resources because of the policy, or to a policy-induced change in preferences, or to both. Disentangling these alternative explanations is essential to develop a theory-consistent demand model and to design effective policies.

Second, the tests of the unitary model typically use data from developed countries or from primarily rural areas of developing countries. Hardly any evidence is available from the urban areas of developing countries. This evidence is greatly needed, as it is important to understand the behavior of the urban poor in a rapidly urbanizing world. The behavior of poor, urban consumers may differ from that of their rural counterparts, if nothing else because they have much less access to food production opportunities.

This paper exploits a policy that makes cash transfers to women. This policy, therefore, exogenously increases both the total income of the targeted households and the share of resources controlled by women. We compare the estimates of the policy impact on eligible households' consumption and budget shares with the predictions from a standard demand model, an Engel curve that relates expenditure shares to total expenditure. The intuition is as follows. We first estimate the policy impacts on food and high-protein food consumption (meat, poultry, eggs, fish, milk,

[1] See, among others, Schultz (1990); Thomas (1990); Bourguignon et al. (1993); Browning et al. (1994); Thomas and Chen (1994); Hoddinott and Haddad (1995); Strauss and Thomas (1995); Udry (1996); Lundberg, Pollak, and Wales (1997); Duflo and Udry (2003); Attanasio and Lechene (2002); Thomas, Contreras, and Frankenberg (2002); Quisumbing and Maluccio (2003); Rangel and Thomas (2005); and Akresh (2008).

and dairy products) using the quasi-experimental variation caused by the introduction of the policy in some areas but not in others. We then estimate the structural parameters of the Engel curves *without* exploiting the policy change. That is, we estimate the budget shares of these two commodities as a function of log consumption and other covariates for two groups of households: those eligible for the program observed before the policy change occurs, and other households with similar characteristics to the first group, but which live in areas without the policy change. We can then use the Engel curve estimates (and the estimates of the impact on total consumption) to predict the changes in budget shares. If the model we estimated accurately reflects individual behavior, these latter changes should not be different from the impacts estimated using the evaluation techniques. If they are different, then, our next step is to figure out why the demand model we have estimated fails.

Our data, a sample of poor urban households in Mexico, were collected to evaluate the impact of a large conditional cash transfer (CCT) program, Oportunidades, in urban areas. Oportunidades provides cash transfers to poor families conditional on making health investments in young children and their mothers and on sending older children to school. The program recipients are the mothers. To be eligible for the transfers, these women have to attend classes on nutrition and health. CCT programs, therefore, can change household demand in two ways: (i) by changing both total household income and the relative incomes of the household members, and (ii) by changing nutrition and health knowledge and preferences.

We choose these data for two reasons. First, we can study the consumption of the urban poor, an under-studied, yet policy-relevant population. Second, the policy variation we use is important to study in itself, because Oportunidades is the prototypical CCT program: it was the first CCT program to be launched on a national scale and to be subjected to a rigorous evaluation. Deemed a success, it is now the flagship welfare program of the Mexican government and the main such program by a large margin, as it covers more than one in ten Mexicans. Because of this success, CCT programs were set up in a large number of countries both within and outside Latin America. However, since its initial evaluation was undertaken using data on rural households, much less is known about its impact on urban households and more broadly about the impact of CCT programs in urban areas.

Unlike in rural areas, the evaluation of the program in urban areas was not based on its random allocation across localities. Rather, "treatment" city blocks, where the program was offered since 2003, were matched to "control" blocks on the basis of several observed characteristics, where the program would be phased-in later. Data were then collected in treatment and control blocks before and after the start of the program.

We begin our empirical analysis by estimating the impact of Oportunidades on food consumption and high-protein food consumption using difference-in-differences matching estimators. We focus on these outcomes for several reasons. First, food is a key outcome of interest, as it accounts for about two-thirds of nondurable consumption for the sampled households, as typically happens for the poor. Second, one of the explicit goals of the program is to improve nutrition and therefore what happens to food consumption and expenditure is particularly important. Third, as the nutritional intakes of the poor are often unbalanced, favoring starchy food such as rice and corn (mainly in rural areas) or highly processed food with little nutritional content (mainly in urban areas), what happens to the consumption of foods rich in proteins is of interest. This is especially true as the diet of the urban Mexicans has shifted away from high-protein food and towards food rich in sugars and refined carbohydrates and sodas (Rivera et al. 2002). Fourth, the evidence from other CCT programs has consistently shown a misalignment between the prediction of standard consumer theory—which, if food is a necessity, implies that an increase in total consumption is associated with a decrease in the share of food consumption, and the empirical evidence—which does not show a decline in food shares.[2] We therefore want to check whether this type of impact is also observed in our context, and whether a standard demand model can predict the empirical evidence. Lastly, since the program impact on high-protein food consumption is large (and estimated relatively precisely), it is particularly interesting to check whether the demand system can predict such a change.

We then estimate three sets of Engel curves using data unaffected by the program. These are eligible households in treatment blocks before the start of the program, and eligible households in control blocks before and after the start of the program in treatment blocks, but before the program was extended to control blocks. We find that food is a necessity and high-protein food a luxury. Armed with these estimates of the demand curves and with the impact of the program on total consumption, we can predict the impacts of the program on expenditure shares and compare them to the quasi-experimental impacts. On the basis of this comparison, we strongly reject the specification of our model. The main reason for this finding, mechanically, is that in the case of the share of food, the quasi-experimental evidence shows an increase, while the Engel curve, showing food as a necessity, predicts a decline. In the case of high-protein food, both the quasi-experimental evidence and the Engel curve predict an increase, but it is much larger in the former than in the latter.

Pointing out that the standard demand model is not an adequate representation of the behavior of our sample of households has important policy implications, since these demand estimates are typically used to predict the impacts of prospective policies. In our case, for example, the use of these estimates to predict policy impacts might lead to a suboptimal provision of similar policies in the future, since the demand curves underestimate the impact of the program on food consumption and high-protein food consumption.

Given our rejection of the basic model, we check whether the program changes some of its parameters. In a sense, our exercise recognizes that the basic model is not "structural" with respect to the policy change in the sense of Lucas (1976), as its parameters move with it. To confirm this conjecture, we estimate the same Engel curves using data collected in 2004, *after* the program was started. We find that in control areas the estimated parameters are similar to the ones from 2002, but in treatment areas they change significantly (both statistically and in magnitude).

[2] The evidence on food is consistent with Maluccio and Flores (2004) for Nicaragua; Paxson and Schady (2007) for Ecuador; Attanasio and Lechene (2011) for rural Mexico; and Attanasio, Battistin, and Mesnard (2011) for Colombia among others.

As we already mentioned, there are at least two explanations for this discrepancy between the estimates of the demand models and of the average treatment effects. One is that the demand curve is based on the erroneous unitary household model, which assumes either homogeneous preferences of the household members or having one decision maker in the household. If members' preferences differ and decisions, which are taken jointly, depend on relative incomes (a proxy for bargaining power), the program might change demand by increasing the female's relative income. The other explanation is that the program changes nutrition knowledge and preferences.

To distinguish between these two possibilities, we reestimate our model on the subset of households headed by a single woman who earned most of the household income regardless of the program transfers and show that, for this group, the parameters of the model do not change with the program. This constitutes indirect evidence that the program changes the mechanisms through which resources are allocated within the household.

In sum, our paper contributes to the literature in several ways. First, it provides evidence on household demand for the urban poor in Mexico. As a by-product of this exercise, it also establishes the impact of the program on food consumption and high-protein food consumption. Second, besides testing and rejecting the unitary model using variation which is arguably exogenous, it provides indirect evidence on the possible reason for this rejection. Third, it generalizes tests of the unitary model à la Browning and Chiappori (1998) in both empirical and conceptual ways: empirically, by estimating more flexible Engel curves; conceptually, by estimating reduced-form parameters without assuming that Oportunidades changes the distribution factors (as we will clarify in the next section). Lastly, it uses policy changes to test structural models. Papers that are relevant for this approach include Lise, Seitz, and Smith (2005); Todd and Wolpin (2006); and Attanasio, Meghir, and Santiago (2011). However, these papers rely on experimental evidence, which we do not have for the urban poor in Mexico. Unlike these papers, therefore, we compare a structural approach with a quasi-experimental one.

I. Structural Models of Consumption Composition

Engel curves, one of the first economic relationships to be analyzed empirically, relate expenditure shares to a total amount of consumption, for a given vector of relative prices.[3] One can derive a demand system that relates the expenditure shares of different commodities to total expenditure, prices, and taste shifters from the optimization problem of an individual or a unitary family that maximizes the utility derived from a set of commodities, given the total amount spent. Indeed, a model in which one assumes that a unitary household maximizes utility, given a budget constraint, implies stringent restrictions on the demand system.

[3] In what follows, we will be using expenditure and consumption interchangeably. This approach is justified by the fact that we focus on nondurable commodities and by the fact that, for urban households, consumption of nonpurchased food is relatively rare.

One widely used system is Deaton and Muellbauer's (1980) Almost Ideal Demand System (AIDS), in which the share of consumption of commodity j in total consumption for household i, w_i^j, is given by the following expression:

$$(1) \qquad w_i^j = \alpha^j(x_i) + \sum_{k=1}^{N} \psi_k^j \ln(p_k) + \theta^j \ln \frac{c_i}{a(\mathbf{p})} + u_i^j.$$

The variable c_i is total consumption and $\mathbf{p}$ a vector of N prices whose generic element is p_k, the price of the k*th* commodity. $a(\mathbf{p})$ is a price index homogeneous of degree 1, which depends on the ψ_k^j and α^j parameters. One can allow the expenditure share to vary by family characteristics x, letting either the intercepts α^j and/or the price coefficients ψ_k^j depend on them. The term u_i^j represents unobserved taste shocks. With appropriate restrictions on the parameters, the shares determined by equation (1) are theory-consistent, in that there exists a well-defined indirect utility function that gives rise to these shares. In other words, if there is a single agent and that agent is endowed with a certain utility function, one can get an equation like (1). The literature refers to such a context, characterized by a unique decision maker, as the unitary model.[4]

If the unitary model is not a good description of the decision process that determines expenditure shares, expenditure shares will not necessarily be described by equation (1). And even if a similar equation was a decent approximation of a set of data in a given context, its parameters would not be "structural" in the sense of Lucas (1976) and could change in response to policy changes, especially those that would induce changes in the relative power of decision makers within the households. We describe a situation where this can happen below.

As it turns out, the unitary model has been rejected in many empirical studies, some of which we have cited in the introduction. Obviously, the demand function's details that better describe actual expenditure shares will depend on the specific model of intra-household allocation one believes to be relevant.[5] A model that has received a considerable amount of attention in the literature, partly because of its tractability, is the so-called collective model, where the only assumption that is made regarding the behavior of a household with multiple decision-makers is that resources are not wasted and are allocated efficiently.

Efficiency is usually defined as a situation in which expenditures are determined so as to maximize the weighted average of two different utility functions, where the weights depend on a number of factors that shift the weight in favor of one or the other individual and that are usually referred to as "distribution factors." Chiappori (1988); Browning and Chiappori (1998); and Bourguignon, Browning, and Chiappori (2009) derive some important results for this case on the way in

[4] Sometimes an equation like (1) is estimated separately for different demographic groups (as defined, for instance, by family composition), therefore allowing for greater flexibility in the way in which these variables affect demand. The unitary model, however, imposes restrictions on the coefficients of equation (1), which guarantee homogeneity, adding-up, and Slutsky symmetry, for each stratum on which the equation is estimated.

[5] Leuthold (1968); Becker (1973); Manser and Brown (1980); McElroy and Horney (1981); Bourguignon (1984); Grossbard-Shechtman (1984); McElroy (1990); Lundberg and Pollak (1993); and Chen and Woolley (2001); among others, propose different types of models of intra-household allocation.

which distribution factors enter the demand system (through a single index) and on the properties of the Slutzky matrix of price elasticities.

Adding distribution factors to share equations such as (1) and testing for the significance of their coefficients has been widely used as a test of the unitary model. Within developing countries, some papers use the introduction of CCT programs to see whether budget shares are affected by changes in distribution factors, on the grounds that the transfers are handed out to women only. For example, Attanasio and Lechene (2002) use the data for the evaluation of PROGRESA, in which the treatment is randomly assigned at the village level, and find that women's income share, instrumented with the treatment assignment dummy, is an important determinant of demand.[6]

The empirical literature on estimating demand systems under the collective model is in its infancy and there are still unresolved issues, both conceptual and empirical. At a conceptual level, it is not easy to derive the form of household demand functions under the collective model, even when one is willing to make some assumptions about the nature of individual utility functions. For instance, in a model with public goods (whose consumption affects the utility of both spouses) even if utility functions are such that in a unitary model demand functions are of the AIDS type, it is not necessarily the case that household demands are AIDS in the collective model.

At an empirical level, most of the papers using the introduction of CCTs as a proxy or instrument for a change in distribution factors overlook the fact that CCTs change knowledge too. As we explain below, in order to receive the transfer, eligible women must attend some classes on nutrition and health. If this knowledge shock has a direct effect on demand, which is not unlikely, the CCT program might change household behavior through both a change in distribution factors and a change in preferences (which may or may not interact with expenditures). In this case, using program eligibility as an instrument for a spouse's income share would be incorrect, as the instrument would be nonexcludable.

We address these shortcomings as follows. We follow Browning and Chiappori (1998) and use an AIDS structure to approximate household demand. However, we let the parameters of the demand system be a function of time, t, (before and after the introduction of the CCT) and space, z, (whether the households live in areas where the program is offered or not). We can therefore rewrite equation (1) as

$$w_i^j = f(z) + \tilde{\alpha}^j(tz)x_i + \tilde{\theta}^j(tz)\ln(c_i) + u_i^j. \tag{2}$$

Notice that equation (2) does not depend explicitly on prices. This omission reflects the fact that the Oportunidades dataset does not contain information on prices. As a consequence, we are unable to estimate price elasticities. However, we control for possible price effects by allowing for state fixed effects.

If estimated on data from control households that do not receive the program, or on *pre-program* data, it may well be that equation (2) fits the data well and, as such, is indistinguishable from an Engel curve derived under the unitary model. Notice

[6] Djebbari (2005); Rubalcava, Teruel, and Thomas (2009); and Attanasio and Lechene (2011) provide additional evidence using the same data. Attanasio, Battistin, and Mesnard (2011) look at rural Colombia; Paxson and Schady (2007) at rural Ecuador; and Maluccio and Flores (2004) at rural Nicaragua.

however that, if program eligibility (and participation) constitutes a distribution factor as it changes the share of resources controlled by women, the parameters of such an equation will not be stable when estimated on households who receive the Oportunidades grant. In this sense, as we mentioned above, the parameters of a version of equation (2) that ignores the distribution factors are not structural.

Equation (2) is more general than the one estimated by Browning and Chiappori (1998) in two ways. First, we do not assume that only changes in distribution factors, in this case the introduction of Oportunidades, may change the shape of the demand curve. That is, changes in the curve parameters over time and space may not necessarily imply a rejection of the unitary model. Second, our curve has flexible intercepts and slopes (by time and space), while the curves estimated by Browning and Chiappori may only have an intercept shift (although the intercept is different for each stratum they consider).

II. *Oportunidades* and Its Evaluation

As mentioned above, in this paper, we study the urban component of Oportunidades, the Mexican conditional cash transfer program. In this section, we provide some information on the nature of the program, on its evaluation sample, and some descriptive statistics on the data we will be using.

A. *Structure, Eligibility, and Benefits of the Program*

The main idea behind Oportunidades, and many other CCTs, is to offer cash (rather than in-kind) transfers to poor households, but imposing some conditions that are intended to stimulate the accumulation of human capital and, therefore, pose the basis for the long run elimination of poverty. All grants are targeted to women. This is important because it can change the balance of power within the household by changing the control of resources for an important income source. Moreover, the receipt of the grant is conditional on a number of activities related to health. The beneficiary mother is supposed to attend some courses and meetings and, if she has young children, to take them to health centers with a certain frequency. These health-related activities can potentially change preferences by giving the beneficiaries information on nutrition and diet. However, in conversations, program officials often report that the courses are not very effective and often fail to engage the beneficiaries.

The Oportunidades grant is also conditional on school enrollment and attendance. Children, starting with the third year of primary school, receive a grant conditional on regular attendance to school. The grant is increasing in the grade attended and, after primary school, when the grant becomes substantially larger than in primary school, a gender differential is also introduced, with girls receiving more than boys.

The targeting of the program is done in several stages. First, the program targets specific geographic areas. Then, within these areas, it targets individual households. During its first phase, between 1997 and 2002, the program was developed mainly in rural Mexico, although some localities were excluded either because they were "too marginal" or because they were not poor enough. In 2002, the government decided to expand the program to urban areas, excluding, however, large metropolitan areas with more than 1 million inhabitants. The urban areas (a definition that is usually larger than a municipality) were chosen on the basis of the prevalence of poverty in the 2000 census. Areas with the highest concentrations of poor households entered the program first, in 2002. We discuss this issue further when describing the evaluation design and strategy.

The individual targeting is based on a score that depends on a number of wealth indicators. While the way in which eligibility is determined in rural and urban areas is essentially the same (households' scores are compared with a cut-off level above which a household is declared eligible for the program), the way in which individual households are approached and registered in the program is substantially different. In rural areas, a census is conducted in each locality and the data collected are used to compute scores and determine the eligibility of each household. In urban areas, instead, the program sets up a local office in a poor neighborhood and publicizes the program, inviting poor households to register for it. When an individual goes to the office, she is administered a questionnaire and preliminary eligibility is established. If the individual is deemed eligible, program officials are dispatched to her residence, where a full survey is administered and, upon confirmation of eligibility, the household is registered with the program.

An implication of the procedure in urban areas is that, to enter the program, individuals first have to approach the local office, without knowing for certain their eligibility status. We suspect that this uncertainty is one cause of the low program take-up rate, which was a little higher than 50 percent of eligible households in 2003, after the first grants were paid. The low participation rate contrasts with very high participation rates in rural areas, where the recipients were informed ex ante that they were eligible for the program grants. An additional likely reason for the low participation rate is that the monetary incentives are much lower in urban than rural areas, as the wages of working teenagers are two to three times as large as the scholarships (Angelucci and Attanasio 2009). Lack of information about the program's existence and features does not seem to be a sizable determinant of the low take-up rate, as this rate remains around 50 percent also in 2004, when the program had been ongoing in urban areas for almost two years.

B. *Evaluation Design and Related Issues*

The program officials first established the city blocks in which the program would expand in 2002. The treatment is not random, as the program first started in the areas with the highest concentration of poor households, as per the information in the 2000 census. This implies that treatment blocks are different from control blocks.[7] The evaluation advisory group sampled blocks within treated areas and matched each block in the treatment sample to a control block with similar characteristics, based on a pre-estimated propensity score. However, the program design prevents one from using certain variables to form the propensity score. For instance, as the program was assigned to the areas with the highest concentrations of poor households, using such a variable would give no intersection between treatment and control samples.

[7] We use the terms "treatment blocks" or "treatment areas" and "control blocks" or "control areas" to indicate in city blocks where the program is implemented in 2003 and 2004 (treatment) or not (control).

Given the sample of treatment and control blocks, a number of individual households were sampled in both groups. These included both eligible households and a small number of ineligible households. In treatment areas, since the number of eligible households participating in the program was lower than expected, it was decided to oversample participants, also including households in blocks adjacent to the treatment blocks. The treatment sample, therefore, is choice-based and the fraction of eligible households participating in the program observed in our treatment sample is quite different from the true fraction of program participants. Fortunately, we can estimate the true proportion of participants in each block from a different dataset, which we use to create weights to make the sample representative of the underlying population. We discuss this issue in Section IIIB.

C. *Sample and Descriptive Statistics*

Our initial sample consists of data on 9,945 eligible households from 267 treatment and 272 control city blocks. Respondents were interviewed in 2002, after households had registered for Oportunidades but before any payments had been made, and in 2004, after payments started in treatment blocks. The data provide information on the consumption of 37 types of food in the week prior to the interview. For non-food items, the survey collects information on the expenditure in the previous month, quarter, or year for different commodities. We transform all the figures to monthly equivalents and they are all expressed in 2002 pesos.

From this initial sample, we have 9,304 valid consumption observations in 2002 and between 7,802 and 7,954 non-missing consumption observations in 2004. The drop in size in 2002 is mainly because of households with incomplete responses. Conversely, in 2004 about 45 percent of the households with missing data have disappeared entirely from the sample. The remaining 55 percent have incomplete responses in the food module. The attrition rates are about the same in control and treatment areas.

To deal with missing observations in the expenditure data, we first regress an indicator for having at least one missing expenditure variable on household income and poverty level, a dummy for missing income, an indicator for living in a treatment area, and year dummies. The results, available upon request, indicate that none of these variables is a statistically significant correlate of having missing expenditure data. Therefore, we keep as many observations as possible for each variable of interest. For example, if a household reports food expenditure but has missing non-food expenditure data, we keep the food observation in our dataset. This results in each variable having a slightly different number of valid observations. The advantage of this procedure is that, by having a larger sample, we increase the precision of the estimates. As a robustness check, we re-estimated all the key parameters, dropping all households with at least one missing observation: the sample size dropped, affecting the standard errors, but the point estimates were largely unchanged.[8]

[8] Because of the outcome of this robustness check and the lack of correlation between having missing consumption data and the household's income and poverty, we chose not to use alternative ways of dealing with missing data, such as multiple imputation.

We also trim the data to account for measurement error. Since our key outcome is the change in log consumption, we trim the top and bottom percentiles of our first-differenced log consumption measures. We do that because, when true consumption is either underreported or overreported, its difference is either too large or too small (the data we trimmed are implausibly high, showing changes in consumption and the two subcategories we consider of almost 200 percent to almost 700 percent). Trimming the data reduces the estimates of Average Intention to Treat parameters by 2 to 3 percentage points.

The numbers of observation used to estimate the 2004 treatment effects are 6,908 for total consumption, 7,341 for food consumption, and 7,132 for high-protein food consumption. The sample size for total consumption is lower because some households do not have non-food consumption data. The sample size is lower for high-protein food than for food because some households did not consume any protein-rich food the previous week in either the 2002 or 2004 data. This prevents us from computing the change in logs. Since the protein-rich food consumption of these households is lower than average, this omission likely biases the estimate of the treatment effect for high-protein food. To address this issue, we estimate this treatment effect a second time after increasing high-protein food consumption by one. Note, however, that we can still estimate the high-protein food share for these households, so reporting zero consumption in high-protein food consumption has no implication for the comparison of the treatment effect estimates and the estimates from the Engel curves. The numbers valid observations for the Engel curves vary from 8,796 to 9,146 in 2002, and from 7,523 to 7,825 in 2004.

Table 1 shows the means of eligible households' socioeconomic variables measured at baseline or earlier in treatment blocks and control blocks and reports the p-values of their difference (with standard errors clustered at the city block level). As expected, the two sets of households are different. Households in treatment blocks have fewer members, are more likely to be headed by a single female, have heads and spouses with slightly worse education, are more likely to have suffered some shocks, and have a lower wealth index. All these features are consistent with them residing in poorer areas. However, neither income nor debt is statistically different for these two groups, and savings are almost two-thirds higher (a statistically significant difference) in treatment areas.

Table 2 reports eligible households' expenditure averages and budget shares for food and high-protein food in treatment and control blocks for 2002 and 2004, as well as the p-values of the differences. Several considerations are in order. First, the households in our sample are, indeed, quite poor. Their consumption is very low, at less than 1,950 and 1,600 pesos per month (roughly \$195 and \$160 in US dollars) in control and treatment blocks, and food represents a very large share of total consumption, accounting for two-thirds of nondurable expenditures. The average expenditure share is not too dissimilar from that observed in rural PROGRESA data. Second, average consumption at baseline is lower in treatment than in control blocks, consistent with the fact that the poorest communities received the program first (although only non-food consumption is statistically different in the two sets of blocks). To estimate impacts, it will be key to take this difference into account. Third, while food consumption decreases and non-food consumption is stable in

Table 1—Comparison of Means of Socioeconomic Variables for Eligible Households in Control and Treatment Blocks, 2002

	Control blocks		Treatment blocks		Difference
	Mean	SD	Mean	SD	[p-value]
Household size 1	0.02	0.14	0.03	0.17	[0.003]***
Household size 2	0.04	0.20	0.07	0.25	[0.000]***
Household size 3	0.08	0.27	0.13	0.34	[0.000]***
Household size 4	0.20	0.40	0.22	0.41	[0.154]
Household size 5	0.23	0.42	0.21	0.41	[0.049]**
Number of kids 0–5	0.82	0.94	0.76	0.91	[0.028]**
Number of kids 6–12	1.16	1.17	0.98	1.10	[0.000]***
Number of kids 13–15	0.32	0.58	0.30	0.58	[0.191]
Number of kids 16–20	0.31	0.65	0.28	0.62	[0.067]*
Number in school 6–12	1.06	1.13	0.90	1.06	[0.000]***
Number in school 13–15	0.22	0.49	0.22	0.49	[0.814]
Number in school 16–20	0.09	0.33	0.09	0.33	[0.701]
Child went to doctor	0.22	0.42	0.22	0.42	[0.967]
Single female	0.19	0.41	0.26	0.45	[0.000]***
No spouse	0.18	0.39	0.25	0.43	[0.000]***
Household head:					
Incomplete primary	0.29	0.45	0.31	0.46	[0.057]*
Complete primary	0.24	0.42	0.20	0.40	[0.000]***
Incomplete secondary	0.05	0.22	0.06	0.23	[0.215]
Complete secondary	0.16	0.37	0.12	0.33	[0.000]***
Higher education	0.05	0.22	0.05	0.22	[0.915]
Employed	0.66	0.48	0.67	0.47	[0.323]
Self-employed	0.20	0.40	0.20	0.40	[0.689]
Went to doctor	0.09	0.28	0.10	0.30	[0.072]*
Spouse:					
Incomplete primary	0.23	0.42	0.24	0.43	[0.284]
Complete primary	0.22	0.42	0.18	0.39	[0.000]***
Incomplete secondary	0.04	0.19	0.04	0.19	[0.744]
Complete secondary	0.13	0.33	0.10	0.30	[0.007]***
Higher education	0.03	0.17	0.03	0.17	[0.873]
Employed	0.14	0.35	0.16	0.37	[0.011]**
Self-employed	0.08	0.27	0.12	0.33	[0.000]***
Went to doctor	0.09	0.29	0.11	0.31	[0.094]*
Income 2002	3,225	20,870	2,839	5,539	[0.287]
Income 2001	696,222	2,688,695	664,562	2,630,553	[0.676]
Income 2000	1,075,103	3,282,003	957,232	3,114,688	[0.267]
Income 1999	1,432,830	3,721,618	1,347,078	3,624,643	[0.507]
Head or spouse worked 2001	0.87	0.33	0.89	0.31	[0.064]*
Head or spouse worked 2000	0.86	0.35	0.87	0.34	[0.253]
Head or spouse worked 1999	0.84	0.36	0.86	0.35	[0.076]*
Savings	48,988	735,842	78,588	930,871	[0.099]*
Debt	79,959	936,692	83,141	953,976	[0.891]
Any death	0.11	0.31	0.12	0.33	[0.057]*
Any job loss	0.10	0.30	0.17	0.37	[0.000]***
Any business loss	0.01	0.07	0.00	0.07	[0.585]
Any natural disaster	0.02	0.15	0.05	0.21	[0.008]***
Wealth	1.65	0.73	1.58	0.71	[0.040]**

Notes: All variables measured at 2002 values, unless otherwise specified. Values of income, savings, and debt at 2002 prices. The last column reports the p-values of t-tests of equality of means in the control and treatment blocks. The null hypothesis is that the difference in means is zero. The standard errors of the differences in means are clustered at the city block level. 9,937 observations.

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

TABLE 2—COMPARISON OF MEANS OF OUTCOME VARIABLES FOR ELIGIBLE HOUSEHOLDS IN CONTROL AND TREATMENT BLOCKS, 2002 AND 2004

	Control blocks		Treatment blocks		Difference
	Mean	SD	Mean	SD	[p-value]
Food consumption 2002	1,942	15,132	1,576	12,770	[0.295]
Non-food consumption 2002	829	737	660	593	[0.000]***
Food/consumption 2002	0.65	0.16	0.66	0.15	[0.105]
High-protein food/consumption 2002	0.19	0.10	0.19	0.10	[0.810]
High-protein food/food consumption 2002	0.30	0.14	0.29	0.14	[0.358]
Food consumption 2004	1,437	841	2,628	22,542	[0.123]
Non-food consumption 2004	822	703	860	1,899	[0.501]
Food/consumption 2004	0.66	0.15	0.67	0.15	[0.367]
High-protein food/consumption 2004	0.20	0.10	0.21	0.10	[0.000]***
High-protein food/food consumption 2004	0.30	0.13	0.32	0.12	[0.000]***

Notes: Values of consumption at 2002 prices. The last column reports the p-values of t-tests of equality of means in the control and treatment blocks. The null hypothesis is that the difference in means is zero. The standard errors of the differences in means are clustered at the city block level. The sample size varies from 7,713 to 9,415 observations depending on which variables we consider.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

control blocks between 2002 and 2004, they both increase considerably in treatment blocks. Finally, we notice that the share of high-protein food, for instance, is higher in treatment than in control communities in 2004.

The data provide information also on a sample of 3,528 ineligible households living in treatment and control blocks. These households are not eligible for the program because their wealth score is too high, i.e., they are not sufficiently poor to be eligible for Oportunidades.

III. The Impact of Oportunidades on Urban Consumption

In this section, we discuss the identification and estimation of Average Intention to Treat effects—that is, the effect of the program on eligible households—and report the estimates of these parameters on the level and logs of total, food, and high-protein food consumption. Our aim is to provide convincing evidence that the assumptions under which we identify and estimate our treatment effects are valid. If this is the case, we can then proceed to compare these quasi-experimental estimates with the Engel curve estimates.

A. *Identification*

Since poverty rates differ systematically in treatment and control areas, we need to control for unobserved determinants of consumption that differ by area to obtain credible impact estimates. We do this by using difference-in-differences matching estimators. First differencing deals with time-invariant unobserved factors, while matching rebalances the sample to deal with time-varying unobserved factors.

Define blocks where the program is offered to eligible households $(Z = 1)$ as "treatment blocks" and blocks where the program is not implemented $(Z = 0)$ as "control blocks." We observe outcomes for households in both block types at time t_1, almost two years after the implementation of Oportunidades, and at time t_0, prior to the program start. Potential outcomes for household i at time t_1 are $Y_{it_1}(1)$ for eligible households living in treatment blocks, $Z_{it_1} = 1$, and $Y_{it_1}(0)$ for eligible households living in control blocks, $Z_{it_1} = 0$. The relationship between potential and observed outcomes is $Y_{it_1} = Y_{it_1}(1)Z_{it_1} + Y_{it_1}(0)(1 - Z_{it_1})$.

Given this notation, the following equation defines the Average Intention to Treat (AIT) effect:

$$AIT = E[Y_{it_1}(1) - Y_{it_1}(0) \,|\, Z_{it_1}(1) = 1].$$

This notation implicitly assumes that potential outcomes for each subject are not affected by the treatment status of others, an assumption usually referred to in the literature as the Stable Unit Treatment Value Assumption (SUTVA), formalized by Rubin (1980, 1986). Our key identification assumption is that, conditional on a set of observed characteristics measured in a preprogram time period $t = t_0, X_{it_0}$, area of residence is independent of the change in potential outcomes $\Delta Y_{it}(1) = Y_{it_1}(1) - Y_{it_0}(1)$ and $\Delta Y_{it}(0) = Y_{it_1}(0) - Y_{it_0}(0)$, i.e., $Z_i \perp \Delta Y_{it}(0), \Delta Y_{it}(1) \,|\, X_{it_0}$. That is, we allow residents of treatment and control blocks to have different levels of potential outcomes, but the differences are assumed to be time invariant, therefore they disappear by taking their first difference.

From the above assumptions, and dropping the subscripts for expositional ease, it follows that $E[\Delta Y(0) \,|\, Z = 1, X] = E[\Delta Y(0) \,|\, Z = 0, X]$. That is,

$$\begin{aligned} AIT_{P(X)=p} &= E[Y_{it_1}(1) - Y_{it_1}(0) \,|\, Z_{it_1}(1) = 1, P(X) = p] \\ &= E\Delta[Y(1) \,|\, Z = 1, P(X) = p] - E[\Delta Y(0) \,|\, Z = 0, P(X) = p], \end{aligned}$$

where we express this parameter as a function of the propensity score $P(X) = P(Z = 1 \,|\, X)$ (Rosenbaum and Rubin 1983). If we further assume common support, i.e., $P(Z = 1 \,|\, X) < 1$, the AIT is

$$AIT = \int_p AIT_{P(X)=p} \, dF(p \,|\, Z = 1).$$

This parameter is identified under the assumptions that the program has no effect in control areas, that the changes in potential consumption in treatment and control areas are independent of areas of residence, conditional on the observed variables, and that there is full common support, $P(Z = 1|X) < 1$.

It is important to provide indirect evidence in support of our identification assumption, which we do below. Before discussing the assumptions' validity, we need to discuss how we deal with our choice-based sampling design.

B. *Dealing with a Choice-Based Sample*

To sample eligible households, the evaluation team used a poverty index built using socioeconomic data from a census of all residents of the selected blocks, collected at baseline. This index was used to select households from control areas that would have been eligible for the program, had the program been implemented in such areas. The sample selection in treatment areas, however, used both the poverty index and administrative data on program enrollment, oversampling participants. The observed fraction of eligible households enrolled in Oportunidades, therefore, is considerably higher than the true one.

We create weights following Manski and Lerman (1977), who show how the weights are the ratio between the true (population) and observed (sample) proportion. We use the baseline census together with the administrative data on program participation to compute the true fraction of eligible households enrolled in Oportunidades by treatment block, Q. We use the sample to compute the observed fraction, H. For each block, the weight for eligible participants is Q/H, while the weight for eligible nonparticipants is $(1 - Q)/(1 - H)$. The weight for eligible households in control blocks is 1, because those households were sampled at random (see Angelucci, Attanasio, and Shaw 2005 for further details).

To estimate our parameters of interest, therefore, we need to use two sets of weights: the Manski and Lerman (1977) weights, to account for the oversampling of eligible participants from treatment blocks, and the matching weights, to rebalance the control areas.[9]

C. *Do the Identification Assumptions Make Sense?*

The consumption of eligible households in control blocks is unlikely to be affected by the program, given that most control and treatment blocks are located in different states, often not geographically contiguous, and that, therefore, control and treatment blocks are distant from each other (Angelucci and Attanasio 2009).

The presence of common support between treatment and control samples is a testable assumption; therefore we proceed to see whether it is maintained in our data. We follow Angelucci and Attanasio (2009) to estimate the propensity score at the individual level by probit using the following variables (and 2002 values, unless otherwise specified): household size dummies; number of children by age categories (0–5, 6–12, 13–15, and 16–20) grouped according to their status (going to school or not); wealth index as a second-order polynomial (program eligibility is based on this index); income (as a second-order polynomial); savings (excluding those of domestic helpers and their relatives, and of individuals whose relationship to other family members is missing) and debt; transitory shocks in 2002 such as death or illness of nonresident family member, job or business loss for resident family member, and whether the household suffered a natural disaster; doctor visits in the previous four weeks for children, head, and spouse (as three separate dummies); household head's and spouse's presence (including multiple heads), gender education dummies (the categories are: no qualification, incomplete primary, complete primary,

[9] To compute standard errors we use the following bootstrap algorithm. First, we sample blocks at random. Second, we estimate the propensity score by probit using the choice-based weights. Third, we use the *psmatch2* code in Stata to generate a counterfactual potential outcome for each treatment observation with the control ones. Fourth, we compute the difference between observed and estimated counterfactual outcome for each treatment observation and estimate the AIT by weighted average using the choice-based weights.

incomplete secondary, complete secondary, higher education), employment status in 2002 (employee or self-employed, the excluded category is unemployed); dummies for whether either head or spouse worked in 1999, 2000, and 2001; income of head and partner in 1999, 2000, and 2001 (as a linear term); and state GDP growth for 2000, 2001, and 2002. These variables are meant to capture systematic differences between treatment and control blocks before the program started.

We show the coefficients of the propensity score in Table 3. These coefficients, estimated using Manski and Lerman (1977) weights, confirm that treatment blocks are poorer than control blocks, as the households living in treatment blocks have lower wealth, a larger share of uneducated household heads, and higher likelihoods of suffering from transitory shocks (except loss of business) and of being headed by females without a partner, normally associated with high indigence. Interestingly, though, residents of treatment blocks also have higher employment rates (both as employees and self-employed) than control block residents, and their income does not differ from the income of control block residents (with the exception of 2001 income, which is higher in treatment blocks), conditional on the other observed characteristics and higher education for the spouse of the household head. Lastly, treatment and control blocks have different state GDP growth rates, confirming they are not balanced at the geographic level. In sum, this table shows the need to rebalance the observed variables between treatment and control blocks.

Figure 1 shows that the common support is complete; that is, for each household in an Oportunidades block, we have a sufficiently high number of close matches from control blocks. Full common support ensures that we can compute average treatment effects for the entire sample of eligible and treated households, respectively, and not only for nonrandom subgroups of families.

We now provide indirect evidence in favor of our conditional independence assumption (CIA). While not directly testable, the evidence provided below supports our conjecture that the CIA holds given the chosen set of conditioning variables.

The main issue for the CIA validity is whether we have successfully controlled for differential trends between treatment and control blocks, since our difference-in-differences approach controls for time-invariant unobserved differences. The fact that control and treatment city blocks are located in different states may be problematic, as these states have different rates of GDP growth. Our first piece of evidence in favor of the CIA validity justifies the need to control for state-specific variables and suggests that, while there are differential trends in income, they disappear once we condition on state GDP growth. Angelucci and Attanasio (2009) show that, while there are differential trends in income between treatment and control blocks, this difference disappears after conditioning on GDP changes (there are no preprogram consumption data, so one cannot compare preprogram trends in consumption).

Our second piece of evidence confirms that adding state GDP growth to the set of variables we use to estimate the propensity score has a sizable effect on the estimated treatment effects. We show this by estimating average treatment effects on the change in log consumption for ineligible households alternatively adding and omitting preprogram state GDP growth. Since these households are not eligible for the program, we expect the treatment effect to be zero. This is exactly what we find

Table 3—Probit Estimates of the Propensity Score—Marginal Effects

	$P(Z = 1 \mid X)$		$P(Z = 1 \mid X)$
Number of kids 0–5	0.017 [0.020]	Spouse employed	0.107 [0.027]***
Number of kids 6–12	−0.008 [0.027]	Spouse self-employed	0.117 [0.031]***
Number of kids 13–15	−0.011 [0.022]	Hh income	0.006 [0.004]
Number of kids 16–20	−0.024 [0.016]	(Hh income)^2	−0.00001 [0.00001]
Number in school 0–5	0.004 [0.020]	Hh income 2001	0.003 [0.001]**
Number in school 6–12	0.0004 [0.021]	Hh income 2000	−0.001 [0.003]
Number in school 13–15	0.002 [0.027]	Hh income 1999	−0.001 [0.001]
Number in school 16–20	−0.021 [0.027]	Head or spouse worked 2001	0.010 [0.031]
Single female	0.122 [0.034]***	Head or spouse worked 2000	0.005 [0.032]
No spouse	0.041 [0.040]	Head or spouse worked 1999	0.067 [0.034]**
Head incomplete primary	−0.055 [0.032]*	Any savings	0.001 [0.006]
Head complete primary	−0.107 [0.039]***	Any debt	0.002 [0.002]
Head incomplete secondary	−0.049 [0.044]	Any death	0.050 [0.027]*
Head complete secondary	−0.115 [0.055]**	Any job loss	0.127 [0.032]***
Head higher education	−0.179 [0.077]**	Any business loss	−0.193 [0.089]**
Spouse incomplete primary	0.136 [0.037]***	Any natural disaster	0.176 [0.067]***
Spouse complete primary	0.12 [0.039]***	Wealth	−0.137 [0.045]***
Spouse incomplete secondary	0.17 [0.043]***	(Wealth)^2	0.006 [0.011]
Spouse complete secondary	0.123 [0.037]***	State GDP growth 2000	−14.37 [2.800]***
Spouse higher education	0.173 [0.043]***	State GDP growth 2001	0.26 [3.706]
Head employed	0.082 [0.041]**	State GDP growth 2002	8.581 [3.196]***
Head self-employed	0.073 [0.042]*		
Area characteristics		No	
Household size dummies		Yes	
Doctor visit dummies		Yes	
Income joint significance		14.84***	
Observations		7,336	

Notes: Marginal effects computed at the mean of the independent variables. Robust standard errors in brackets; clustering at the city block level. Hh = household. Omitted education category: uneducated. Omitted employment category: unemployed. Unless otherwise specified, all variables are from 2002. State GDP growth (0.01 = 1 percent; from INEGI's Banco de Informacion Economica, http://dgcnesyp.inegi.gob.mx/cgi-win/bdieintsi.exe/Consultar.

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

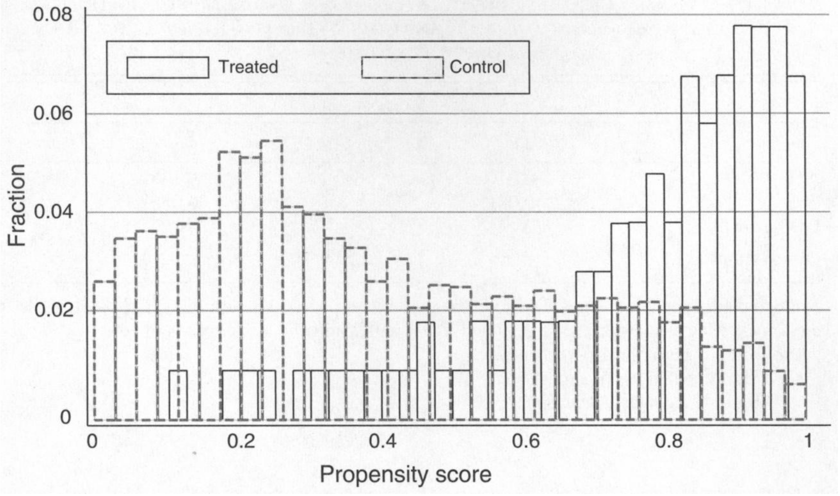

Figure 1. The Propensity Score

Note: To compute this propensity score we use the estimates from Table 3.

when we condition on GDP growth: Table 4 shows that the effect of Oportunidades on ineligible households' log consumption is −0.010 and not statistically significant (column 1). However, when we fail to control for the difference in GDP growth, we estimate a positive, statistically significant, and large treatment effect: consumption appears to be over 14 percent higher for ineligible households in treatment areas (column 2).[10]

D. *Estimates of the Average Intention to Treat Effects*

We estimate the 2004 AIT effects for total expenditure, food, and high-protein food (in logs). These are the key parameters of interest for the purpose of our exercise, because we will use them to make a comparison with the predictions from the Engel curves, as we discuss later. These parameters are estimated from the same population for which we will estimate the Engel curves, using the same weighting scheme.

We employ a difference-in-differences local linear regression matching estimator with a tricube kernel. Since neither plug-in nor cross-validation bandwidth selectors seem to perform well for finite samples (Frolich 2005), we tried many different bandwidths using log food consumption as our outcome. The counterfactual mean of the change in log consumption is roughly stable for bandwidths between 0.1 and 0.3 and between 0.6 and 0.9 (which means we use centered subsets of $N \times$ bandwidth observations, where NS is the number of observations). The difference

[10] We explore these issues in greater detail in Angelucci and Attanasio (2009). Among other things, we show that using different geographic controls—for example, state dummies—as an alternative to GDP growth prevents us from finding good matches for at least 60 percent of the households in the treatment blocks. We believe that using state-level GDP growth to estimate the propensity score strikes the right balance between controlling for differential trends between control and treatment areas and having full common support.

TABLE 4—AVERAGE TREATMENT EFFECT (ATE) ON LOG CONSUMPTION FOR INELIGIBLE HOUSEHOLDS CONTROLLING (*column 1*) AND NOT CONTROLLING (*column 2*) FOR GDP GROWTH BY STATE

	Log consumption	
	(1)	(2)
ATE	−0.010	0.148
	[0.081]	[0.056]***
GDP growth	Yes	No
Observations	3,528	

Notes: Standard errors estimated with the block-bootstrap (1,000 repetitions). The block is the city block. Matching estimates using local linear regression and a bandwidth of 0.6.
***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

between the various estimates is at most 1 percentage point.[11] We find a similar pattern when the outcome is the log of protein-rich food consumption. However, the choice of bandwidth matters more when the outcome is log consumption. We report estimates of the AIT effects using a tricube kernel and bandwidths of 0.1 and 0.6. The smaller bandwidth is our preferred one because it reduces the bias (but increases the variance) of the estimates. We provide the results from the larger bandwidth for comparative purposes.

We conduct balancing tests on the matched sample, in the spirit of Smith and Todd (2005). As pointed out in Lee (2006), one should not use rigid rules in interpreting the results from these tests and rather focus on whether the differences between the treatment and control means of the propensity score covariates become smaller after matching. As such, we point out that, although the means of several rebalanced variables are statistically different, the pseudo-R^2 of the probit estimate of the propensity score drops from 0.29 for the unbalanced sample to 0.02 for the rebalanced one. Moreover, the average absolute bias decreases by one half after the rebalancing. Based on this evidence, we consider the rebalanced control group to be similar enough to our treatment group (at least in terms of the considered covariates).[12,13]

We compute the standard errors using the block bootstrap to allow for area-specific shocks; the block is the city block. We estimate the propensity score each time we resample the data and present estimates of the standard errors using 1,000 repetitions.

[11] We obtained a very similar counterfactual mean to the one with local linear regression, a tricube kernel, and a bandwidth of 0.1 using three nearest neighbors with replacement and propensity score inverse weighting.

[12] We experimented extensively both with changing the functional form of the covariates, e.g., using higher order polynomials in income, and with changing the set of covariates itself. In no case did we find a combination of different functional forms or a different set of covariates that further reduces the bias and the pseudo-R^2. The detailed results of these tests are available upon request.

[13] We use the Stata *pstest* command to perform the balancing tests. Since this code cannot perform balancing tests with a tricube kernel—which is computationally much faster to use in the estimation of the AIT than any other kernel—we perform these tests using a biweight kernel. This kernel weights the observations in a similar way to the tricube. The results discussed above are obtained using a bandwidth of 0.1. However, the results from the balancing test are minimally affected by the bandwidth choice. We thank Barbara Sianesi for making this suggestion.

Table 5—Average Intention to Treat Estimates for 2004 Log Total, Food, and High-Protein Food Consumption

	Bandwidth = 0.1 (1)	Bandwidth = 0.6 (2)
Log total consumption	0.041 [0.034]	0.020 [0.035]
Observations	6,908	6,908
Log food consumption	0.071 [0.030]***	0.056 [0.029]**
Observations	7,341	7,341
Log high-protein food consumption	0.174 [0.040]***	0.168 [0.039]***
Observations	7,132	7,132
Log high-protein food consumption[a]	0.240 [0.050]***	0.225 [0.045]***
Observations	7,341	7,341

Notes: Block-bootstrap standard errors in brackets, the block is the city block (1,000 repetitions). Local linear regression matching estimates with a tricube kernel and bandwidths of 0.1 and 0.6. We took the difference in log consumption (and subcategories) between 2004 and 2002, trimming the top and bottom percentile.

[a] Estimated increasing all high-protein food consumption by 1.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

The key finding is that the program increases food consumption in percentage terms more than total consumption (although this difference is not statistically significant at conventional levels), as can be seen by comparing the first two rows of Table 5. This result differs quantitatively from the effect of CCT programs in rural Mexico (Attanasio and Lechene 2011) and in other countries' rural areas, e.g., in Colombia (Attanasio, Battistin, and Mesnard 2011), where the budget share of food consumption has no statistically significant increase and is roughly constant. Schady and Rosero (2008), however, find that in rural Ecuador, a CCT program increases the share of food in total expenditure.

Overall, the AIT effect of the program on consumption is about 4 percent and that on food consumption is about 7 percent, when we use a bandwidth of 0.1, and 2 percent and 5.6 percent with a bandwidth of 0.6. The last two rows of Table 5 show a large AIT effect on high-protein food consumption. These estimates are stable to the bandwidth choice and vary from about 17 percent to 24 percent, depending on how we treat households reporting zero consumption in at least one data wave. Note that the inclusion of this latter group of households increases the point estimates, suggesting that households with zero consumption are the ones for which the program causes the highest proportional increase. The total log consumption AITs are not statistically different from zero at conventional levels, unlike the other two AITs, which are much more precisely estimated.

The fact that, in many CCT programs, the share of food in consumption does not decrease is at odds with the notion that food is a necessity and, as such, should increase *less* than proportionally with total consumption. Some potential explanations of this fact are that food consumption might not be a necessity for some of the households in our sample. Moreover, we should worry about the different level of data quality for food and non-food items. It is possible that the program changes the process of resource allocation within the household and, consequently, the household's demand, by giving cash transfers to women. To address these issues, we turn to the estimation of Engel curves; that is, to the relationship between expenditure shares and total expenditure.

IV. Engel Curves of Mexican Urban Poor

Before specifying and estimating Engel curves for food and high-protein food, we need to specify which measures of "total consumption" we use in the various cases. As is well known, one can use two-stage budgeting arguments to focus on a specific subproblem of the problem faced by a consumer (or a household) in determining the allocation of resources among different commodities. When specifying the Engel curve for food, we will consider as total expenditure the amount spent on nondurable goods and services. This approach takes the total expenditure in a period as given and considers its allocation among different commodities. The standard two-stage budgeting approach takes into account the allocation of resources over time. If resources other than nondurable consumption and services are consumed in a given period, this approach implicitly assumes that these other consumption items (such as the services from durables) and nondurables are additively separable.

Analogously, one can model the consumption of high-protein food as a function of either total nondurable consumption or total food expenditure. In the latter case, one needs to assume that food consumption is separable from the rest of consumption. This assumption allows us to avoid possible measurement problems with the non-food items. Such an assumption might not be too far-fetched in our context, as food constitutes such an important part of the total budget.

We estimate the Engel curves for food and high-protein food (meat, fish, eggs, and dairy products). For the latter variable, we consider the relationship between expenditures on protein-rich food and food consumption only. We estimate four sets of Engel curves for both food and high-protein food: before and after the start of the program—that is, in 2002 and 2004—and for households potentially eligible for the program who live in treatment and control areas.

Treatment and control areas have different characteristics, as they were not chosen through a randomization. This motivates us to estimate the Engel curves separately for treatment and control areas even for the period before the introduction of the program. Comparing the evolution of the Engel curve shapes estimated in control and treatment areas before and after the introduction of the program is the main purpose of this exercise.

We estimate the relationship between the share of food $(j = f)$ or high-protein food $(j = p)$ and the log of total expenditure $(\ln c^f)$ or food expenditure $(\ln c^p)$ separately for households (i) in treatment $(T = 1)$ and control $(T = 0)$ areas and for 2002 $(t = 2)$ and 2004 $(t = 4)$:

$$w_{it}^{jT} = \gamma_{0t}^{jT} + \gamma_{1t}^{jT} \ln (c_{it}^{jT}) + \gamma_{2t}^{jT} x_i^T + u_{it}^{jT}. \tag{3}$$

We control for several predetermined demographic and economic variables, x_i. These variables, measured at 2002 levels, are a set of dummies for household size, single- and female-headed households, head and spouse literacy, head employment or self-employment, and whether the household suffered any of the following shocks in 2002: death, unemployment, business loss of a household member, and natural disaster. We also add the following variables: number of children in the 2002 age groups 0–5, 6–12, 13–15, and 16–20; number of children attending school in 2002 for the previous age groups (except the 0–5 one); a household wealth score; and state dummies. This is the set of characteristics we use throughout. The state dummies are important to capture the possibility of differences in relative prices (which are unobserved in our data) in different areas.[14] We also experimented with municipal dummies, obtaining similar results (available upon request).

In unreported regressions, we estimated this relationship both nonparametrically and semiparametrically, without imposing a functional form relationship for log expenditure or log food expenditure. The estimated curves are approximately linear. We also estimated equation (3) using a second-order polynomial in log expenditure. The coefficient of the quadratic term was never statistically different from zero. Therefore, we focus on estimating the linear specification above.[15]

Total or food consumption are likely endogenous in equation (2) because of both unobserved taste shifters that could affect both total consumption and its allocation and measurement error in consumption. Measurement error is potentially problematic because consumption appears both in the denominator of the dependent variable and in logarithmic form as an explanatory variable.[16] We follow the literature by using log average hourly local wages as well as log monthly household income as instruments for expenditures (adding a dummy for households with missing income data). For each specification, we present both OLS and IV estimates, where the two expenditure variables are instrumented with log total income, log municipality wages, and a dummy for missing income. We notice that instrumenting is important: the IV and OLS estimates of the consumption coefficients are statistically different from each other.

Table 6 shows the parametric OLS and IV estimates of the food Engel curves as a function of total consumption for control and treatment areas (columns 1–2 and 3–4) and for 2002 and 2004 (top and bottom panel). The OLS and IV estimates of the parameters of interest tend to be different from each other, especially in control areas. The Engel curve is downward sloping in both types of areas, consistent with the notion that food is a necessity, and the slope is steeper in control areas (-0.1 versus -0.05). Between 2002 and 2004, the curve becomes steeper in control areas

[14] Note the trade-off between using conditioning variables measured at a different point in time from the dependent variable, potentially estimating a different parameter, and using contemporaneous values of the covariates, which means possibly conditioning on a variable affected by the treatment.

[15] See Banks, Blundell, and Lewbel (1997) on the potential importance of having a second-order polynomial in log expenditure.

[16] Since this form of measurement error is not classical, the IV estimator is inconsistent and Lewbel (1996) suggests using a different estimator. However, Attanasio, Battistin, and Mesnard (2011) show that there is no fundamental difference between the estimates from the Lewbel (1996) and the IV estimators. They used data from a conditional cash transfer program in Colombia similar to Oportunidades. This program targeted households with food expenditure shares similar to those in our sample.

TABLE 6—ENGEL CURVES FOR FOOD CONSUMPTION
(*log consumption on the right-hand side*)

	Control blocks		Treatment blocks	
	OLS (1)	IV (2)	OLS (3)	IV (4)
Panel A. 2002				
Log consumption	−0.034 [0.009]***	−0.101 [0.022]***	−0.031 [0.008]***	−0.055 [0.021]***
Constant	0.907 [0.071]***	1.356 [0.170]***	0.885 [0.058]***	1.071 [0.165]***
Observations	3,271	3,271	5,743	5,743
First stage IV significance (*F*-stat)		26.20		53.67
Panel B. 2004				
Log consumption	−0.069 [0.007]***	−0.112 [0.015]***	−0.045 [0.019]**	−0.043 [0.019]**
Constant	1.205 [0.054]***	1.443 [0.115]***	1.008 [0.144]***	0.992 [0.148]***
Observations	2,839	2,837	4,874	4,874
First stage IV significance (*F*-stat)		75.29		74.81

Notes: Standard errors in brackets, clustered at the city block level. The covariates are listed after equation (3) in the text. Excluded instruments are log total income, log municipality wages, and a dummy for missing income. The coefficients of the X variables may differ for different household groups.

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

and flatter in treatment areas. To test whether these changes are statistically significant, Figure 2 plots the estimated curves for a specific x_i^T vector of average and modal household characteristics. We use the average wealth and the modal state of residence and demographics of households in treatment areas: these are the states of Colima, Chiapas, and Sinaloa, having an employed household head, a literate head and spouse, two children (aged 6–12 and 13–15 and both in school), and a household size of 4. The modal household experiences no shocks.[17] The shape and position of the Engel curve for control areas is not statistically different between 2002 and 2004. Conversely, the food Engel curve in treatment areas in 2004, which is higher and flatter than in 2002 for the selected household type, is significantly different from the 2002 estimated curve. We should stress that in all specifications we control for state dummies that are supposed to capture, among other factors, the effect that relative prices might have on the expenditure shares.

Table 7 shows the parametric OLS and IV estimates of the high-protein food Engel curves as a function of food consumption. We implicitly assume separability between food and non-food consumption. In 2002, the share of high-protein food is an increasing function of food consumption in both treatment and control blocks.

[17] We have dummies for groups of adjacent states, and not for individual states, because a large group of treatment and control areas are sampled from different states.

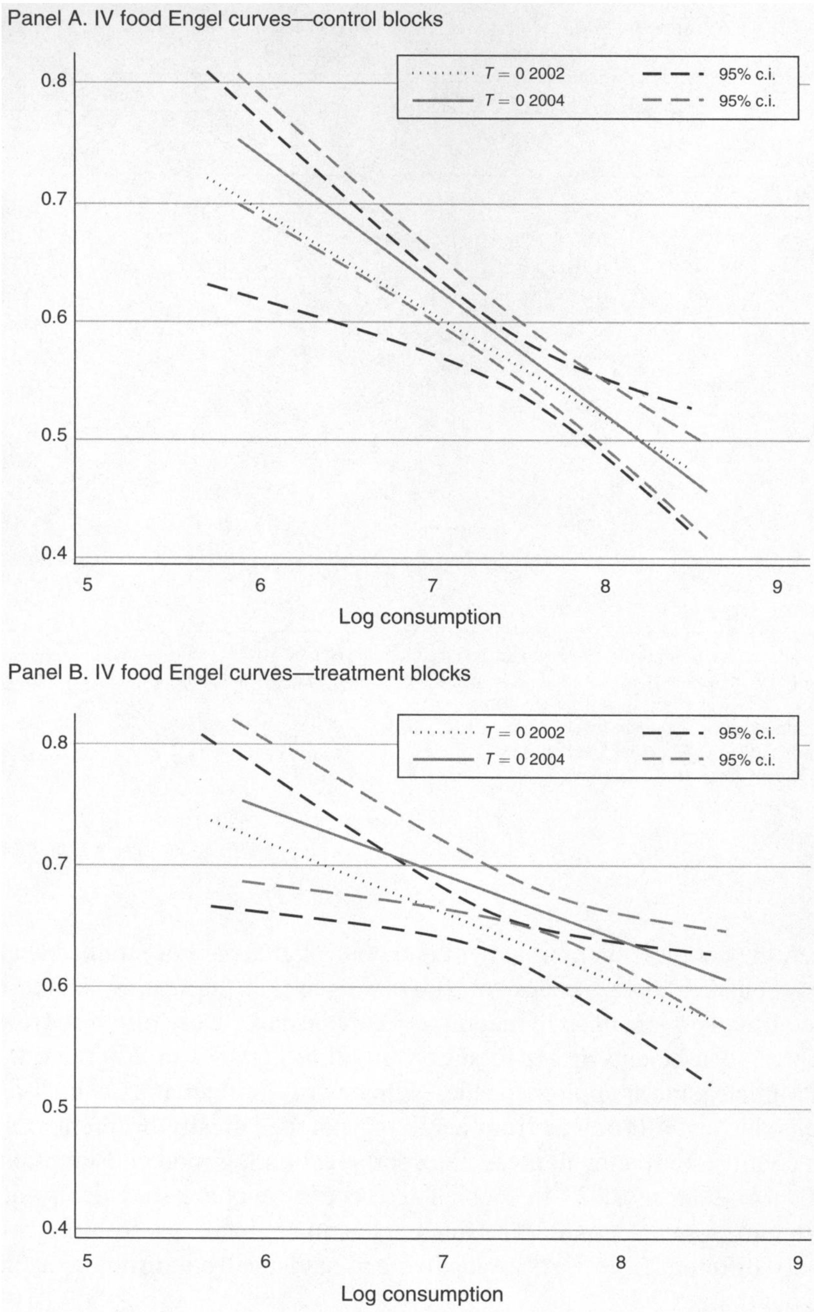

FIGURE 2. FOOD ENGEL CURVES (LOG CONSUMPTION), 2002 AND 2004

Note: These lines are drawn using the estimates from Table 8 (column 2 for panel A and column 4 for panel B).

TABLE 7—ENGEL CURVES FOR HIGH-PROTEIN FOOD CONSUMPTION
(*log food consumption on the right-hand side*)

	Control blocks		Treatment blocks	
	OLS (1)	IV (2)	OLS (3)	IV (4)
Panel A. 2002				
Log food consumption	0.018	0.073	0.033	0.094
	[0.007]**	[0.025]***	[0.007]***	[0.018]***
Constant	0.193	−0.295	0.073	−0.375
	[0.055]***	[0.177]*	[0.051]	[0.134]***
Observations	3,376	3,376	5,919	5,919
First stage IV significance (F-stat)		17.88		38.41
Panel B. 2004				
Log food consumption	0.038	0.082	0.021	0.037
	[0.008]***	[0.016]***	[0.015]	[0.014]***
Constant	0.048	−0.315	0.180	0.062
	[0.062]	[0.111]***	[0.104]*	[0.100]
Observations	2,917	2,915	5,032	5,032
First stage IV significance (F-stat)		51.03		65.97

Notes: Standard errors in brackets, clustered at the city block level. The covariates are listed after equation (3) in the text. Excluded instruments are log total income, log municipality wages, and a dummy for missing income. The coefficients of the X variables may differ for different household groups.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Moreover, the coefficients in treatment and control areas are of similar magnitude. When we consider the estimates for 2004, we see that the results for the control areas are virtually unchanged in magnitude and not statistically different from those for 2002: the coefficient on log food consumption is 0.073 in 2002 and 0.082 in 2004. A similar pattern applies for the estimates of the intercepts, which are both negative, statistically different from zero, and not statistically different from each other. In contrast, in treatment areas, the coefficient on log food consumption drops from 0.094 in 2002 to 0.037 in 2004. These coefficients are statistically different from zero and from each other. The intercept estimate increases from negative and statistically different from zero to positive and statistically indistinguishable from zero.

These changes are reflected in Figure 3 where the shape and position of the Engel curve for control areas is not statistically different between 2002 and 2004, while in treatment areas the Engel curve for high-protein food is shifted up and is considerably flatter. In sum, high-protein food seems less of a luxury than before the program.

While there might be different reasons behind these findings, one possibility is that the program, by targeting transfers to women, changes the process of resource allocation within the family and therefore the Engel curves.

V. Treatment Effects and Engel Curves

The Engel curves represent a simple structural model that can be used to predict the impact of the program on the structure of consumption. Under the assumption that the difference-in-differences matching estimates of the AITs are consistent, we can use those estimates to validate the structural model. That is, we can compare the predictions from the Engel curves with the difference-in-differences matching estimates of the AITs that we estimated in Section III.

Suppose, for instance, that expenditure shares, such as the share of food or high-protein food, are determined by the following equation:

$$w_{it}^{j} = \theta_0^j + \theta_1^j \ln(c_{it}) + u_{it}, \tag{4}$$

where the index j refers to either food (f) or high-protein food (p), the indices i and t to the household and time (we have omitted demographics and other control variables for notational simplicity). As before, the variable c_{it} is food expenditure. The following equivalence is true:

$$AIT_{w^j} \equiv \Delta w_{it}^{j} = \theta_1^j \Delta \ln(x_{it}) \equiv \theta_1^j AIT_{\ln x}, \tag{5}$$

where the parameters AIT_{w^j} and $AIT_{\ln x}$ are the Average Intention to Treat effects on budget shares and log consumption for food shares (or log food consumption for high-protein food shares). This equivalence suggests that if we multiply the estimate of the program impact on $\ln(x_{it})$ from Section III, $AIT_{\ln x}$, by the estimate of the Engel curve coefficient of log expenditure, θ_1^j, we should obtain an estimate of the Average Intent to Treat on budget shares, AIT_{w^j}. That is, if both the difference-in-differences matching estimates and the Engel curve estimates are consistent, the two sets of estimates of AIT_{w^j} should not be statistically different.

In the previous section, however, we saw that for each of our two outcomes of interest (food shares and high-protein food shares) we have four different estimates of the Engel curve parameters relevant for the prediction in equation (5), corresponding to the four different cells over which we estimated the parameters of the Engel curve: 2002 and 2004 (that is, before and after the beginning of the program) and treatment and control areas (see Tables 6 and 7). If the household demand theory under which the Engel curve was derived is correct, the four estimates should be the same (or not statistically different), as they all refer to the same underlying structural relationship between log expenditure and budget shares and one could use any of them. Conversely, if the households in our sample behave in a way inconsistent with the unitary model, the estimated Engel curve is misspecified and there is no reason why the estimate of $\theta_1^j\, AIT_{\ln x}$ should be the same as the estimate of AIT_{w^j}.

The fact that the shape of the Engel curves changes in treatment areas after the introduction of the program in 2003 implies that we will obtain different predictions. Table 8 reports our predictions of the impact of the program on the share of food in total expenditure and of the share of high protein food in total food expenditure using each of the estimates we obtained in Tables 6 and 7.

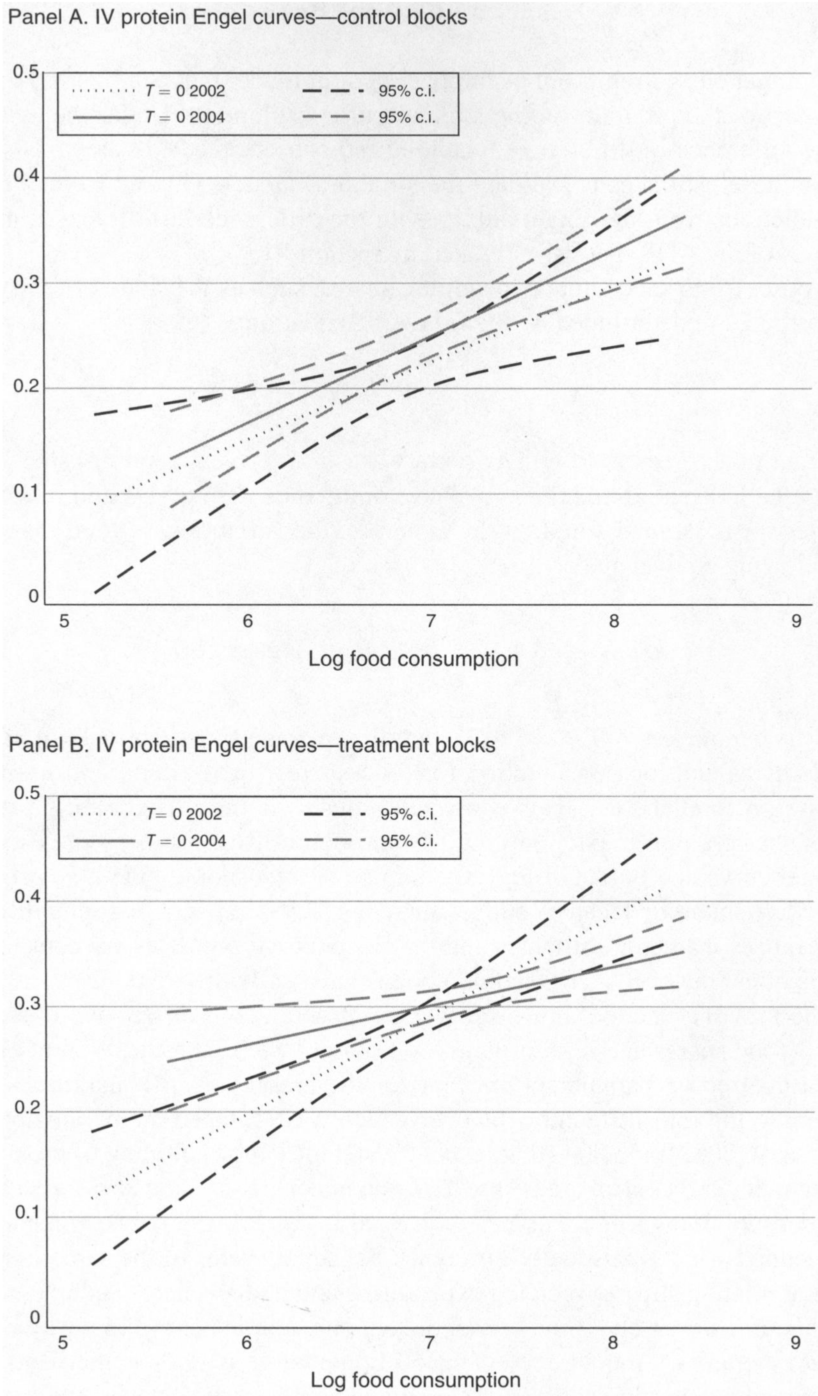

FIGURE 3. HIGH-PROTEIN FOOD ENGEL CURVES (LOG FOOD CONSUMPTION), 2002 AND 2004

Note: These lines are drawn using the estimates from Table 9 (column 2 for panel A and column 4 for panel B).

Table 8—Food and High-Protein Food Shares (w): Treatment Effects (*column 1*) and Estimates from the Engel Curves (*columns 2 to 4*)

	Matching estimates	Predictions from Engel curves		
	AIT_w (1)	$\beta_2^T AIT_{\ln x}$ (2)	$\beta_2^C AIT_{\ln x}$ (3)	$\beta_4^C AIT_{\ln x}$ (4)
Panel A. Bandwidth = 0.1				
Food consumption	0.018	−0.002	−0.004	−0.004
($x = c$: Engel curve with ln c on RHS)	[0.014]	[0.002]	[0.004]	[0.004]
High-protein food consumption	0.041	0.006	0.007	0.006
($x = fc$: Engel curve with ln fc on RHS)	[0.009]***	[0.003]*	[0.004]*	[0.003]**
Panel B. Bandwidth = 0.6				
Food consumption	0.024	−0.001	−0.002	−0.002
($x = c$: Engel curve with ln c on RHS)	[0.014]*	[0.002]	[0.004]	[0.004]
High-protein food consumption	0.039	0.004	0.005	0.005
($x = fc$: Engel curve with ln fc on RHS)	[0.009]***	[0.003]	[0.004]	[0.003]*

Notes: fc = food consumption; c = total nondurable consumption; RHS = right hand side. Standard errors estimated with the block-bootstrap, the block is the city block (1,000 repetitions). Local linear regression matching estimates of the AITs with a tricube kernel and bandwidths of 0.1 and 0.6. IV estimates of the β parameters. Excluded instruments are log total income, log municipality wages, and a dummy for missing income.
***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Table 8 reports the various estimates of the AIT for the budget share of food in total consumption, AIT_{w^f}, and of high-protein food in total food, AIT_{w^p}. We begin by reporting the difference-in-differences matching estimates of this parameter (column 1), estimated under the assumptions and methods discussed in Section III. We then compute the same parameters using the estimated slope from three different Engel curves—the 2002 curve estimated for the treatment area (column 2) and the 2002 and 2004 curves estimated for the control areas (columns 3 and 4)—multiplied by the difference-in-differences matching estimates of $AIT_{\ln x}$ shown in Table 5 (where this parameter is the AIT effect on log total consumption for the food budget share and the AIT effect on log food consumption for the high-protein food budget share). As before, we report two sets of estimates, produced using alternatively bandwidths of 0.1 and 0.6. The results are not sensitive to the bandwidth choice.

Our preferred comparison is the one between the estimates from the first two columns. The two estimates of the same parameters should not be statistically different if (i) the difference-in-differences matching estimates of the treatment effects are consistent and, if (ii) we can interpret the estimate of the slope of the Engel curve as a structural parameter, i.e., as being an accurate representation of the behavior of the households in our sample. For the former, the crucial nontestable hypothesis in this case is that, conditional on observed characteristics, the change in potential outcomes in the absence of the treatment is independent of the area of residence. The tests we performed provide some indirect evidence that this hypothesis holds in our data. We will judge the ability of the Engel curves to predict changes in consumption shares as indirect evidence on the stability of the structural parameters.

We first notice that both for total food shares and for the shares of high-protein food, the estimates from columns 2 to 4 do not appear to be statistically different from each other. However, they are different from the AIT estimates we report in column 1. In the case of the food share, the Engel curves predict a modest decline between 0.2 percent and 0.4 percent. However, the AIT estimates imply an *increase* in the share between 1.8 and 2.4 points. The relatively low precision of these estimates makes the differences only marginally significant. However, the evidence is much more dramatic for the share of high-protein food in total food. For this commodity, the Engel curve implies an increase in the share between 0.4 percent and 0.7 percent (which is at best marginally significantly different from zero). The AIT instead implies a much larger increase, between 3.9 percent and 4.1 percent. These estimates are statistically different from those implied by the structural model and therefore imply a strong rejection of the model.[18] These findings constitute strong evidence that the Engel curve derived from a unitary model is misspecified, in the sense that it is not able to predict the impact of the program in treatment areas.

The failure of the Engel curves to predict the impact of the program is consistent with the apparent structural shift in its parameters between 2002 and 2004 in treatment areas. The issue, of course, is why one would get different parameter estimates and how we interpret these differences.

From a theoretical point of view, if the AIDS structure is an adequate representation within a unitary framework of the allocation of resources within the household, changes in relative prices could induce differences in intercepts but not in the slope of the Engel curve we estimated.[19] In the absence of systematic changes in prices, the intercepts should also be stable.

Our findings that the Engel curve fails to predict the change in food consumption observed after the introduction of the CCT program is similar to the evidence from other CCT programs—e.g., in rural Mexico (Attanasio and Lechene 2002 and 2011); rural Colombia (Attanasio, Battistin, and Mesnard 2011); rural Ecuador (Schady and Rosero 2008); and rural Nicaragua (Maluccio and Flores 2004). As already mentioned, however, there are two alternative explanations for our results. One explanation is that the inconsistency between the prediction of the Engel curve and the estimated effects is caused by handing the transfer to women. This would change the nature of intra-household allocation of resources and increase the bargaining power of women, whose preferences may differ from their husbands', the typical main earners in these households. That is, the misspecification of the Engel curve could arise because of the failure of the unitary model behind the derivation of equation (1). An alternative explanation for our results is that the CCT program changes eligible households' information on nutrition and health by providing access to health care and to nutrition and health classes. In this case, the household still acts in ways consistent with the unitary model and the Engel curve is misspecified because it does not account for the change in preference caused by the CCT program.[20]

[18] The result for high-protein food is robust to the consideration of total consumption (rather than total food consumption) as the "expenditure" argument in the Engel curve. It is therefore not explained by the assumption of separability between food and non-food consumption.

[19] In the case of a QUAIDS system, the presence of quadratic terms implies interactions between total expenditure and prices. As we do not observe prices, price changes could induce changes in the slope of the Engel curve. However, as we mentioned above, we have no evidence of nonlinearities in the Engel curves we have estimated.

[20] The program may also change the preferences of a bargaining household, so these two explanations of our findings are not mutually exclusive.

VI. A Test Using Female-Headed Households

A possible interpretation of the results shown above is that Oportunidades increases eligible wives' bargaining power, resulting in changes in the composition of expenditures not predicted by standard Engel curves. A simple validation test of this hypothesis consists of comparing the change in high-protein food shares obtained from the Engel curves with the AIT estimates of the same parameter for households whose women's bargaining power is *not* increased by Oportunidades. For these households, the Engel curve should be stable over time and the different estimates of the same parameter should not be statistically and qualitatively different from each other.

We identify such households as those in which the female is designated as the head of household, as the husband is not present in 2002, and in which the income of the head accounts for at least 60 percent of total income. We include the second restriction to avoid considering situations in which the mother lives with older adults who might have some decision-making power. A similar exercise was performed by Schady and Rosero (2008) for Ecuador.

Table 9 shows the high-protein food consumption AIT (column 1) estimated with local linear regressions matching, a tricube kernel, and bandwidths of 0.1 and 0.6 (top and bottom panels) for a sample of 746 households. Columns 2 to 4 predict the impact of the program on the share of high-protein food using the estimated Engel curves in the same way as described for Table 8. The size of the program effect on the share of high-protein food consumption varies between 2.4 and 3.5 percentage points and is not statistically different from the size of the effect observed in the whole sample, although its point estimate is slightly lower. However, unlike the results for the whole sample, it is no longer statistically different from the estimates in columns 2 to 4, obtained by multiplying the estimated treatment effect on log food consumption by the estimate of the slope of the Engel curve for single, female-headed households. While it is true that, given the reduced size of the sample, the estimates in Table 9 are not extremely precise, the point estimates in the four columns are of the same order of magnitude, unlike the point estimates in Table 8.

These findings are consistent with our interpretation of the results presented in Section V—namely that the traditional Engel curve is misspecified for the whole sample because the unitary household model is not an adequate representation of household behavior. Establishing that the Engel curves do not change for households in which the female is already the main decision maker discredits the competing explanation that Oportunidades may change preferences for a healthy diet by either transmitting knowledge about nutrition or providing preventative health care.

VII. Conclusions

This paper uses the policy change caused by the introduction of Oportunidades, Mexico's flagship welfare program, to study the demand for food and for high-protein food among poor urban households eligible for this program. We model their demand using a theory-consistent, state-of-the-art demand system and estimate it paying due care to a number of methodological and econometric issues. We

TABLE 9—HIGH-PROTEIN FOOD SHARES (w):
TREATMENT EFFECTS (*column 1*) AND ESTIMATES FROM THE ENGEL CURVES (*columns 2 to 4*) FOR SINGLE, FEMALE-HEADED HOUSEHOLDS WITH HIGH EX ANTE BARGAINING POWER

	Matching estimates	Predictions from Engel curves		
	AIT_w (1)	$\beta_2^T AIT_{\ln fc}$ (2)	$\beta_2^C AIT_{\ln fc}$ (3)	$\beta_4^C AIT_{\ln fc}$ (4)
Panel A. Bandwidth = 0.1				
High-protein food consumption	0.035	0.021	0.041	0.013
	[0.025]	[0.018]	[0.033]	[0.017]
Panel B. Bandwidth = 0.6				
High-protein food consumption	0.024	0.021	0.042	0.013
	[0.025]	[0.019]	[0.035]	[0.018]

Notes: fc = food consumption. Standard errors estimated with the block-bootstrap, the block is the city block (1,000 repetitions). Local linear regression matching estimates of the AITs with a tricube kernel and bandwidths of 0.1 and 0.6. IV estimates of the β parameters. Excluded instruments are log total income, log municipality wages, and a dummy for missing income. The sample size is 746. We consider single female household heads who earn at least 60 percent of the household income.
***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

investigate whether eligible households' demand for total food and for high-protein food changes in a way consistent with the prediction from the Engel curves estimated using preprogram consumption.

We find that eligible households consume much more food and, in particular, much more high-protein food than would be predicted by a standard Engel curve, estimated on data from the same population observed before the beginning of the program. The fact that women start to control a sizable proportion of the family income seems to induce a change in the way households allocate total expenditure among different commodities.

These findings, together with others in this literature, call for a new, theory-consistent demand model that does not assume a unitary household. These findings also have important policy implications, especially if the transfers policy-makers are concerned with are targeted to specific economically weak agents in the population, such as women.

REFERENCES

Akresh, Richard. 2008. "(In)Efficiency in Intra-Household Allocations." https://netfiles.uiuc.edu/akresh/www/Research/Akresh_IntraHH.pdf.

Angelucci, Manuela, and Orazio Attanasio. 2009. "Oportunidades: Program Effects on Consumption, Low Participation, and Methodological Issues." *Economic Development and Cultural Change* 57 (3): 479–506.

Angelucci, Manuela, and Orazio Attanasio. 2013. "The Demand for Food of Poor Urban Mexican Households: Understanding Policy Impacts Using Structural Models: Dataset." *American Economic Journal: Economic Policy.* http://dx.doi.org/10.1257/pol.5.1.146.

Angelucci, Manuela, Orazio Attanasio, and Jonathan Shaw. 2005. "The Effect of Oportunidades on the Level and Composition of Consumption in Urban Areas." In *External Evaluation of the Impact of Oportunidades Program 2004: Education*, edited by B. Hernandez-Prado and M. Hernandez-Avila, Vol. 4, 105–52. Mexico City: INSP.

Angelucci, Manuela, and Giacomo De Giorgi. 2009. "Indirect Effects of an Aid Program: How Do Cash Transfers Affect Ineligibles' Consumption?" *American Economic Review* 99 (1): 486–508.

Attanasio, Orazio, Ettore Battistin, and Alice Mesnard. 2011. "Food and Cash Transfers: Evidence from Colombia." *Economic Journal* 122 (559): 92–124.

Attanasio, Orazio, and Valérie Lechene. 2002. "Tests of Income Pooling in Household Decisions." *Review of Economic Dynamics* 5 (4): 720–48.

Attanasio, Orazio, and Valérie Lechene. 2011. "Efficient Responses to Targeted Cash Transfers." http://economics.mit.edu/files/6620.

Attanasio, Orazio, Costas Meghir, and Ana Santiago. 2012. "Education Choices in Mexico: Using a Structural Model and a Randomized Experiment to Evaluate PROGRESA." *Review of Economic Studies* 79 (1): 37–66.

Banks, James, Richard Blundell, and Arthur Lewbel. 1997. "Quadratic Engel Curves and Consumer Demand." *Review of Economics and Statistics* 79 (4): 527–39.

Becker, Gary S. 1973. "A Theory of Marriage: Part I." *Journal of Political Economy* 81 (4): 813–46.

Bourguignon, François. 1984. "Rationalité Individuelle Ou Rationalité Strategique: Le Cas de L'offre Familiale de Travail." *Revue Economique* 35 (1): 147–62.

Bourguignon, François, Martin Browning, and Pierre-André Chiappori. 2009. "Efficient Intra-Household Allocations and Distribution Factors: Implications and Identification." *Review of Economic Studies* 76 (2): 503–28.

Bourguignon, François, Martin Browning, Pierre-André Chiappori, and Valérie Lechene. 1993. "Intra-Household Allocation of Consumption: A Model and Some Evidence from French Data." *Annales d'Economie et de Statistique* (8) 29: 137–56.

Browning, Martin, and Pierre-André Chiappori. 1998. "Efficient Intra-Household Allocations: A General Characterization and Empirical Tests." *Econometrica* 66 (6): 1241–78.

Browning, Martin, François Bourguignon, Pierre-André Chiappori, and Valérie Lechene. 1994. "Income and Outcomes: A Structural Model of Intrahousehold Allocation." *Journal of Political Economy* 102 (6): 1067–96.

Busso, Matias, John Dinardo, and Justin McCrary. 2009a. "Finite Sample Properties of Semiparametric Estimators of Average Treatment Effects." http://emlab.berkeley.edu/~jmccrary/BDM_JBES.pdf.

Busso, Matias, John Dinardo, and Justin McCrary. 2009b. "New Evidence on the Finite Sample Properties of Propensity Score Matching and Reweighting Estimators." http://economics.mit.edu/files/764.

Chen, Zhiqi, and Frances Woolley. 2001. "A Cournot-Nash Model of Family Decision Making." *Economic Journal* 111 (474): 722–48.

Chiappori, Pierre-André. 1988. "Rational Household Labor Supply." *Econometrica* 56 (1): 63–90.

Deaton, Angus S., and John Muellbauer. 1980. "An Almost Ideal Demand System." *American Economic Review* 70 (3): 312–26.

Djebbari, Habiba. 2005. "The Impact on Nutrition of the Intra-Household Distribution of Power." Institute for the Study of Labor (IZA) Discussion Paper 1701.

Duflo, Esther, and Christopher Udry. 2003. "Intra-Household Resource Allocation in Côte d'Ivoire: Social Norms, Separate Accounts, and Consumption Choices." Unpublished.

Frolich, Markus. 2005. "Matching Estimators and Optimal Bandwidth Choice." *Statistics and Computing* 15 (3): 197–215.

Hoddinott, John, and Lawrence Haddad. 1995. "Does Female Income Share Influence Household Expenditures? Evidence from Côte d'Ivoire." *Oxford Bulletin of Economics and Statistics* 57 (1) 77–96.

Hoddinott, John, and Emmanuel Skoufias. 2003. "The Impact of PROGRESA on Food Consumption." International Food Policy Research Institute (IFPR) Food Consumption and Nutrition Division Discussion Paper 150.

Lee, Wang-Sheng. 2006. "Propensity Score Matching and Variations on the Balancing Test." Unpublished.

Leuthold, J. H. 1968. "An Empirical Study of Formula Income Transfers and the Work Decision of the Poor." *Journal of Human Resources* 3 (3): 312–23.

Lewbel, Arthur. 1996. "Demand Estimation with Expenditure Measurement Errors on the Left and Right Hand Side." *Review of Economics and Statistics* 78 (4): 718–25.

Lise, Jeremy, Shannon Seitz, and Jeffrey Smith. 2005. "Evaluating Search and Matching Models Using Experimental Data." Institute for the Study of Labor (IZA) Discussion Paper 1717.

Lucas, Robert E., Jr. 1976. "Econometric Policy Evaluation: A Critique." *Carnegie-Rochester Conference Series on Public Policy* 1 (1): 19–46.

Lundberg, Shelly J., and Robert A. Pollak. 1993. "Separate Spheres Bargaining and the Marriage Market." *Journal of Political Economy* 101 (6): 988–1010.

Lundberg, Shelly J., and Robert A. Pollak. 2003. "Efficiency in Marriage." *Review of Economics of the Household* 1 (3): 153–67.

Lundberg, Shelly J., Robert A. Pollak, and Terence J. Wales. 1997. "Do Husbands and Wives Pool Their Resources? Evidence from the United Kingdom Child Benefit." *Journal of Human Resources* 32 (3): 463–80.

Maluccio, John A., and Rafael Flores. 2004. "Impact Evaluation of a Conditional Cash Transfer Program: The Nicaraguan Red de Protección Social." International Food Policy Research Institute (IFPR) Food Consumption and Nutrition Division (FCND) Discussion Paper 184.

Manser, Marilyn, and Murray Brown. 1980. "Marriage and Household Decision-Making: A Bargaining Analysis." *International Economic Review* 21 (1): 31–44.

Manski, Charles F., and Steven R. Lerman. 1977. "The Estimation of Choice Probabilities from Choice Based Samples." *Econometrica* 45 (8): 1977–88.

McElroy, Marjorie B. 1990. "The Empirical Content of Nash-Bargained Household Behavior." *Journal of Human Resources* 25 (4): 559–83.

McElroy, Marjorie B., and Mary Jean Horney. 1981. "Nash-Bargained Household Decisions: Toward a Generalization of the Theory of Demand." *International Economic Review* 22 (2): 333–49.

Moffitt, Robert A. 1996. "Identification of Causal Effects Using Instrumental Variables: Comment." *Journal of the American Statistical Association* 91 (434): 462–65.

Paxson, Christina, and Norbert Schady. 2007. "Cognitive Development among Young Children in Ecuador: The Roles of Wealth, Health, and Parenting." *Journal of Human Resources* 42 (1): 49–84.

Quisumbing, Agnes R., and John A. Maluccio. 2003. "Resources at Marriage and Intra-Household Allocation: Evidence from Bangladesh, Ethiopia, Indonesia, and South Africa." *Oxford Bulletin of Economics and Statistics* 65 (3): 283–327.

Rangel, Marcos, and Duncan Thomas. 2005. "Out of West Africa: Evidence on the Efficient Allocation of Resources within Farm Households." University of Chicago Harris School of Public Policy Working Paper 05.15.

Rivera, Juan A., Simón Barquera, Fabricio Campirano, Ismael Campos, Margarita Safdie, and Victor Tovar. 2002. "Epidemiological and Nutritional Transition in Mexico: Rapid Increase of Non-Communicable Chronic Diseases and Obesity." *Public Health Nutrition* 5 (1A): 113–22.

Rosenbaum, Paul, and Donald Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70 (1): 41–55.

Rubalcava, Luis, Graciela Teruel, and Duncan Thomas. 2009. "Investments, Time Preferences, and Public Transfers Paid to Women." *Economic Development and Cultural Change* 57 (3): 507–38.

Rubin, Donald. 1980. "Comment." *Journal of the American Statistical Association* 75 (371): 591–93.

Rubin, Donald. 1986. "Comment: Which Ifs Have Causal Answers?" *Journal of the American Statistical Association* 81 (396): 961–62.

Schady, Norbert, and Jose Rosero. 2008. "Are Cash Transfers Made to Women Spent Like Other Sources of Income?" *Economics Letters* 101 (3): 246–48.

Schultz, T. Paul. 1990. "Testing the Neoclassical Model of Family Labor Supply and Fertility." *Journal of Human Resources* 25 (4): 599–634.

Smith, Jeffrey, and Petra Todd. 2005. "Does Matching Overcome LaLonde's Critique of Non-Experimental Estimators?" *Journal of Econometrics* 125 (1–2): 305–53.

Strauss, John, Germano Mwabu, and Kathleen Beegle. 2000. "Intra-Household Allocations: A Review of Theories and Empirical Evidence." *Journal of African Economies* 9 (Supplement 1): 83–143.

Strauss, John, and Duncan Thomas. 1995. "Human Resources: Empirical Modeling of Household and Family Decisions." In *Handbook of Development Economics*, Vol. 3A, edited by T. N. Srinivasan and Jere Behrman, 1883–2023. North Holland: Amsterdam.

Thomas, Duncan. 1990. "Intra-Household Resource Allocation: An Inferential Approach." *Journal of Human Resources* 25 (4): 635–64.

Thomas, Duncan, and Chien-Liang Chen. 1994. "Income Shares and Shares of Income: Empirical Tests of Models of Household Resource Allocations." Research and Development (RAND) Labor and Population Program Working Paper 94-08.

Thomas, Duncan, Dante Contreras, and Elizabeth Frankenberg. 2002. "Distribution of Power Within the Household and Child Health." Unpublished.

Todd, Petra E., and Kenneth I. Wolpin. 2006. "Assessing the Impact of a School Subsidy Program in Mexico: Using a Social Experiment to Validate a Dynamic Behavioral Model of Child Schooling and Fertility." *American Economic Review* 96 (5): 1384–1417.

Udry, Christopher. 1996. "Gender, Agricultural Production, and the Theory of the Household." *Journal of Political Economy* 104 (5): 1010–46.